CHARACTERIZATION OF THE TWO GENES ENCODING CYTOPLASMIC RIBOSOMAL PROTEIN L23A IN *ARABIDOPSIS THALIANA*

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

in the Department of Biology

University of Saskatchewan

Saskatoon

By

Kerri Bryn McIntosh

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Biology
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5E2

ABSTRACT

RPL23a is one of the ~80 ribosomal proteins (r-proteins) of the cytoplasmic ribosome in the model plant *Arabidopsis thaliana*. The objectives of this research were to establish Arabidopsis RPL23a as a functional r-protein, characterize expression patterns for the two genes (*RPL23aA* and *B*) encoding RPL23a using reverse transcription PCR (RT-PCR), and identify regulatory elements controlling the expression of *RPL23aA* and *B*. Complementation of a yeast *l25* mutant with AtRPL23aA demonstrated that AtRPL23aA can fulfill all the essential functions of L25 *in vivo*. A survey of various Arabidopsis tissue types showed that, while *RPL23aA* and *B* expression patterns both showed increased transcript abundance in mitotically active tissues, *RPL23aB* transcript levels were generally lower than those of *RPL23aA* and responded differently to abiotic stresses. In order to determine *cis* regulatory elements controlling *RPL23aA* and *B* expression, the 5' regulatory region (RR) of each gene was characterized via plants carrying a series of 5' RR deletion fragments upstream of a reporter. Transcript start sites and 5' untranslated regions (UTRs) for both *RPL23aA* and *B* were also characterized using primer extension, and transcripts from 5' deletion transgenics were amplified using RT-PCR. No correlation was observed between putative *cis*-acting elements and the expression patterns from the *RPL23aA* and *B* deletion transgenics, although a 102 bp sequence in the *RPL23aB* 5' RR was found to contain pollen-specific elements. 5' leader introns were found in each *RPL23a* gene, and amplification of transgene transcripts from deletion series plants indicated the importance of post-transcriptional and translational regulation in *RPL23aA* and *B* expression. This thesis work is the first demonstration of a plant RPL23a protein as a functional member of the L23/L25 (L23p) conserved r-protein family, and is one of the few in-depth studies of the regulation of r-protein genes in plants. While the majority of previous research on plant r-protein gene expression has focused solely on transcript levels, I show herein that post-transcriptional mechanisms have a critical role in regulating these genes, and thus plant r-protein genes more strongly resemble their mammalian counterparts than those of yeast in terms of structure and regulation.

ACKNOWLEDGEMENTS

First, I would like to thank my supervisor and mentor, Dr. Peta Bonham-Smith, for her guidance and support throughout my graduate career. I will always be grateful for the introduction to that most extraordinary of complexes, the ribosome.

I would also like to thank my committee members, Dr. Arthur Davis, Dr. Susan Kaminskyj, Dr. William Roesler, and the Head of Department Dr. François Messier (and former Heads Dr. Larry Fowke and Dr. Vipen Sawhney) for their continued support and work on my behalf. I appreciate the insight you have contributed to my research throughout my graduate studies career.

I would like to thank the members of the Bonham-Smith lab, past (Laureen Blahut-Beatty, Ashley Byun, Rebecca Cross, Katherine Lockhart, Jeffrey Pylatuik, Lester Young), present (Rory Degenhardt, Jacqueline Hulm, Donna Lindsay), and 'honorary' (Shobha Devi Potlakayala), for their support, guidance, and friendship.

I would also like to thank the faculty, graduate students, and staff (especially Dennis Dyck, Marlynn Mierau, Evelyn Peters, Jeaniene Smith, Joan Virgl, and Deidre Wasyliw) of the Department of Biology for all of the work they have done on my behalf.

Most of all, I would like to thank my family, Linda, Brian, and Carli McIntosh. I cannot possibly thank you enough for your endless support and love, and all the practical help you have given me throughout this often painful process. Thank you for everything. And thank you to Ian Potter, for nine years of love, support, and patience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A site/loop	aminoacyl site/loop
ABA	(±)cis, trans-abscisic acid
BAP	6-benzylaminopurine
bp	base pair
cDNA	complementary DNA
DPA	days post-anthesis
E site	exit site
EF	elongation factor
EJC	exon junction complex
ETS	external transcribed spacer
FR	flanking region
GA_3	gibberellic acid
GUS	β-glucuronidase
hnRNA	heterogeneous nuclear RNA
IAA	indole-3-acetic acid
IF	translation initiation factor
ITS	internal transcribed spacer
LSU	large subunit
mRNA	messenger RNA
NOR	nucleolar organizing region
nt	nucleotide
P site/loop	peptidyl site/loop
pI	isoelectric point
PTC	peptidyl transferase center
RACE	rapid amplification of cDNA ends
RACK	receptor of activated C-kinase
RNP	ribonucleoprotein
r-protein/RP	ribosomal protein
RR	regulatory region
rRNA	ribosomal RNA
RT-PCR	reverse transcription polymerase chain reaction
SR	signal recognition particle receptor
SRD	sarcin-ricin domain
sRNA	soluble RNA (tRNA)
SRP	signal recognition particle
SSU	small subunit
TF	trigger factor
TMS	transmembrane sequence
TOP	terminal oligopyrimidine tract
TOR	target of rapamycin
tRNA	transfer RNA
UTR	untranslated region
wt	wild type

CHAPTER 1. LITERATURE REVIEW

1.1. Introduction

Ribosomes are the ribonucleoprotein bodies responsible for polypeptide synthesis in all living cells. Often identified as organelles, these membrane-free 2.4-4.5 MDa particles are more accurately described as the largest enzymatic complexes in the cell, mRNA-directed polymerases that form peptide bonds between amino acids via a peptidyl transferase function. The ribosome is dynamic in nature; the large and small subunits that comprise each ribosome associate and dissociate as they attach to, or release, messenger RNA (mRNA); nucleic acid and protein components of the subunits shift as translation progresses and individual protein components associate with, and dissociate from, the complex. The process of subunit and, ultimately, ribosome assembly, is itself a highly complex and dynamic process, responding to the ever changing needs of the cell.

Assembly of the eukaryotic cytoplasmic ribosome is a particularly complex process requiring the coordinated production and transport of four ribosomal RNAs (rRNAs) and ~80 ribosomal proteins (r-proteins), in approximately equimolar amounts. While three rRNA genes are arranged in a multi-copy transcription unit and are transcribed in the nucleolus by RNA Polymerase I (Pol I), the 5S (Svedberg/sedimentation coefficient) rRNA and r-protein genes are dispersed throughout the genome and are transcribed by two other polymerases. 5S rRNA genes are transcribed in the nucleus by RNA Pol III and the resulting 5S rRNA is transported to the nucleolus. R-protein genes are transcribed in the nucleus by RNA Pol II, their mRNAs are transported to the cytosol for translation, and the majority of r-proteins are subsequently transported through the nucleus, to the nucleolus, for subunit assembly. Following assembly, the subunits are exported through nuclear pore complexes to the cytosol, where they assemble on mRNA to form complete ribosomes. Ribosome assembly represents one of the most challenging regulatory and transport events within the eukaryotic cell.

Given the complexity of the assembly of the eukaryotic ribosome and its subunits, an important question remains: how is the expression of ribosomal constituents regulated as part of ribosome biogenesis? Little is known about how the expression of such a large number of widely dispersed genes is controlled. Although many nucleic acid and protein components of the ribosome are among the most highly evolutionarily conserved molecules across all kingdoms of life, the ways in which the genes encoding these molecules are regulated differs between species and between the genes themselves. Regulation of r-protein gene expression can occur at the transcriptional, post-transcriptional, translational, and post-translational levels, depending on the organism (reviewed in Mager, 1988; Larson et al., 1991). While ribosome structure, function, biogenesis, and ribosomal gene expression have been studied in detail in organisms such as *Escherichia coli*, *Saccharomyces cerevisiae*, and, to a lesser extent, in mammals such as *Rattus norvegicus*, little is known about the ribosome and its regulation in plants. In many plants, including *Arabidopsis thaliana*, the presence of multiple expressed genes encoding each ribosomal protein further complicates coordinated regulation.

The following thesis research is an investigation of the regulation of a large subunit r-protein gene family (*RPL23a*) in the model plant Arabidopsis. The RPL23a homologue in other organisms has been shown to be a 'primary binder' (binds specifically and independently) of 23S and eukaryotic 23S-like (25S, 26S, or 28S) rRNA (El-Baradi et al., 1984; 1985), playing a key role in the formation of the large ribosomal subunit. This thesis research has established the functional equivalence of AtRPL23a and yeast L25 using a complementation study (McIntosh and Bonham-Smith, 2001; Chapter 2, this volume); compared expression profiles, at the transcriptional level, of the two genes encoding RPL23a (McIntosh and Bonham-Smith, 2005; Chapter 3, this volume); and characterized some of the regulatory mechanisms of the genes encoding RPL23a using transcript start site mapping, reverse transcription PCR (RT-PCR), and the generation of transgenic plants carrying a series of 5' regulatory region deletion fragments upstream of a reporter gene (Chapter 4, this volume).

1.2. Historical Overview

1.2.1. 1781-1940s: Cytology and the biochemistry of RNA-rich particles

Ribosome research began long before ribosomes themselves were identified and named, and has continued to advance as technological development has allowed. The nucleolus, site of the rRNA genes and ribosome subunit assembly, was first documented by Fontana in 1781 (reviewed in Miller, 1981), but the first observations localizing ribosomes in the cell were not made until over a century later (1897) when strands of basophilic cytoplasm (dubbed "ergastoplasm" by Garnier) were identified in mammalian cells (reviewed in Bielka, 1982). Throughout the 1940s, the tools of UV spectrophotometry combined with cytology and use of ribonuclease treatments proved useful in studying the role played by RNA in the metabolism of the cell (reviewed in Bielka, 1982). Experiments investigated RNA content in the nucleolus and the cytoplasm (Caspersson and Schultz, 1940), and correlated high cytoplasmic RNA content with mitotically active plant and animal tissues (Caspersson and Schultz, 1939). The basophilia associated with the cytoplasm of active, dividing cells was thus identified as being a result of high RNA content (Caspersson and Schultz, 1939).

1.2.2. 1940s-1957: Sedimentation fractions and electron microscopy

The invention of the ultracentrifuge by Theodor Svedberg in the early 1920s made it possible to separate cellular components at high speed through a density gradient. Differential centrifugation enabled the identification of the cytoplasmic fraction containing ribonucleoprotein particles and associated phospholipids initially believed to be mitochondrial in nature (Claude, 1941); the ribonucleoprotein-phospholipid particles were later called "microsomes" (Claude, 1943). Electron microscopy (EM), introduced in the 1930s, allowed a much closer look at the contents of the cytoplasm. EM studies of animal (primarily rat liver) cells identified small, ~100-150 Å, 'granules' both associated with the endoplasmic reticulum (ER) and free in the cytoplasm (Palade, 1955). Palade related these particles to the microsomal fraction and to the basophilic ergastoplasm, hypothesizing that it was most likely the small granules that contained RNA, conferring the affinity for basic dyes (Palade, 1955). Further EM examination of rat liver cells and microsome fractions confirmed that the microsomes were composed of the previously identified granules and ER membranes and

ribonuclease treatment identified the small particles as the probable source of microsomal RNA (Palade and Siekevitz, 1956). Isolation of the ribonucleoprotein particles from the microsomal fraction confirmed the RNA content of the particles, which were also, via radiolabeled amino acid incorporation experiments, found to be involved in protein synthesis (Littlefield et al., 1955). It was during the 1950s that the sedimentation coefficients characteristic of purified eukaryotic and prokaryotic ribosomes and their subunits were determined. While these early studies showed a range of coefficients for the ribonucleoproteins they examined (e.g. 74S ribosomes from pea epicotyls, Ts'o et al., 1956; 77.5S rat liver ribosomes, Petermann and Hamilton, 1957), they indicated the approximate sedimentation coefficients recognized as standard today: 80S eukaryotic ribosomes with 60S and 40S subunits, and 70S prokaryotic ribosomes with 50S and 30S subunits (reviewed in Bielka, 1982).

1.2.3. 1958: The microsome gets a new name

The biochemical and cytological discoveries surrounding the ribosome and protein synthesis throughout the 1950s culminated in a 1958 meeting of the Biophysical Society on "microsomal particles and protein synthesis" (Roberts, 1958). It was at this meeting that Richard Roberts coined the term 'ribosome' for the ribonucleoprotein particles (20 to 100S) of the microsomal fraction of the cell (Roberts, 1958). Not only would the designation of 'ribosome' distinguish the protein particles from the lipids of the microsome, but Roberts argued that the term also "has a pleasant sound" (Roberts, 1958). Interestingly, the proportion of research presented at the Biophysical Society symposium that was focused on bacterial (9 of 20 papers) and mammalian (8 of 20 papers) systems reflected the current favored ribosome research subjects; only two papers were presented on yeast ribosomes, and a single paper on plants – a trend which has continued throughout the history of ribosome research up to the present day. Following the symposium, work on the process of translation and the factors involved, including transfer RNA (tRNA; initially identified as soluble RNA/sRNA), messenger RNA (mRNA; template RNA), rRNA, and protein factors, continued into the 1960s. By the early 1960s, it was clear that ribosomes were ubiquitous and performed the same function in all living organisms (Ts'o, 1962).

1.2.4. 1960s: The genetic code and modeling translation

The discovery of the structure of DNA (Watson and Crick, 1953[a,b,c]) opened the door for solving the genetic code, the translation of which was then envisioned as a templating process much like DNA replication (Woese, 2001). The challenge of the code was answered in the 1960s; Nirenberg and Matthaie (1961) determined that template RNA was necessary for peptide synthesis (along with tRNA and ribosomes), and that a synthetic polyuridylic acid (polyU) template would result in production of a phenylalanine chain. This was followed in the mid-1960s by further elucidation of the triplet codes for each amino acid (Nirenberg et al., 1965; Söll et al., 1965). With the triplet code for amino acids resolved, the process of translation, by which polypeptides are synthesized according to the genetic code, remained unclear. In the 1960s, with little detail known about translation or the ribosome itself, the first version of the A (aminoacyl)-site-P (peptidyl)-site (initially identified as the AA-sRNA and protein binding sites) 'cartoon' model of translation was proposed (Watson, 1964). Although this model has continued to be adapted and modified over the last 40 years some favor rejecting this simple cartoon concept of translation altogether in light of today's much more complex structural knowledge of the ribosome (Woese, 2001).

1.2.5. 1970s: Two-dimensional electrophoresis and ribosomal proteins

While the rRNA components of the large and small subunits of the ribosome were identified between 1959 and 1968 in both prokaryotes and eukaryotes (reviewed in Bielka, 1982), it was not until the 1970s that the development of 2-dimensional polyacrylamide gel electrophoresis (2D-PAGE) allowed researchers to visualize the individual proteins of the ribosome for the first time. At this time the proteins of the large and small subunits from *E. coli* (Kaltschmidt and Wittmann, 1970) and rat liver (Sherton and Wool, 1972; Welfle et al., 1972) ribosomes were first fractionated, demonstrating the wide array of ribosomal proteins (r-proteins) present in both subunits. Kaltschmidt and Wittmann (1970) developed a standard 2D gel technique used to fractionate ribosomes and constituent subunits, and they introduced the r-protein naming system now commonly used; i.e. the designation S for small subunit proteins (e.g. S8) or L for large subunit proteins (e.g. L23) followed by a number assigned based on gel position, the higher the molecular weight the lower the number. 2D electrophoresis

studies continued throughout the 1970s, examining primarily bacterial and mammalian ribosomes, although a small amount of work was carried out on plant r-proteins (Gualerzi et al., 1974). The number of different 2D gel patterns obtained using different electrophoresis conditions necessitated the development of a standard r-protein nomenclature using a single set of running conditions. A standard nomenclature for mammalian r-proteins was thus proposed in a collaborative effort based on a variety of ribosomes (rat, rabbit, mouse, HeLa, L, and Chinese hamster ovary/CHO cells) fractionated according to a standard protocol and correlated with previously published numbering systems (McConkey et al., 1979). Along with the fractionation of ribosomes, the stoichiometry of each r-protein within the ribosome had yet to be determined. Hardy (1975), examining the 70S ribosomes of *E. coli*, concluded that all r-proteins but L7 and L12 were present in equimolar amounts, a single copy per ribosome. Although the methods used by Hardy were later disputed (Tal et al., 1990), the conclusion that all *E. coli* r-proteins (except L7 and L12, lateral stalk proteins of the large subunit) are present in a single molecule per ribosome has not been significantly changed (Tal et al., 1990).

1.2.6. 1970s – 1980s: Seeking the peptidyl transferase – RNA or protein?

The characterization of the enzymatic activity of the ribosome has primarily been carried out in *E. coli*. By the late 1960s, work on this model prokaryote had identified the enzymatic activity catalyzing peptide bond formation between amino acids, peptidyl transferase (named by Maden et al., 1968), as a function of the ribosome; the center of catalysis was isolated somewhere on the large ribosomal subunit (Monro, 1967; Maden et al., 1968). The Monro group had developed a ribosome-catalyzed "fragment reaction" between a formylmethionyl-tRNA fragment (*N*-formylmethionyl-CAACCA) and puromycin (mimics the aminoacyl end of an aminoacyl-tRNA) which effectively isolated the peptidyl transferase component of the ribosome without dependence on GTP, template, or any protein factors (Monro, 1967); this reaction became widely used in subsequent peptidyl transferase studies. With the analysis of the individual components of the ribosome there came a concerted effort to localize the peptidyl transferase activity to one or more ribosomal constituents; the assumption was that r-proteins must be involved, as the only enzymes identified at that time were proteins. During the 1970s and into the 1980s, r-proteins were stripped away from the ribosome,

subunits reconstituted, and tRNAs crosslinked to the ribosome in an attempt to localize the peptidyl transferase activity of the *E. coli* ribosome (reviewed in Noller, 1993). The primary tool used for removing r-proteins was LiCl extraction, forming LiCl 50S "cores." While these treatments did succeed in abolishing peptidyl transferase activity, even the most stringent of LiCl treatments could not remove all r-proteins from core particles, leaving behind ~12-20 proteins of the subunit (Hampl et al., 1981). Successive r-protein removal/core constitution experiments determined that L11 (Nierhaus and Montejo, 1973), L16 (Moore et al., 1975), or a combination of ~6 r-proteins (L2, L3, L4, L15, L16, L18) and 23S rRNA were considered 'essential' for re-establishing peptidyl transferase activity to stripped subunits (Hampl et al., 1981). Other r-proteins (e.g. L20, L24) were defined as being important for early subunit assembly or as 'helper' r-proteins, enhancing the enzymatic function of other subunit r-proteins (Hampl et al., 1981). Although it was hypothesized that the r-proteins required to restore peptidyl transferase activity were probably responsible for the actual enzymatic function, at least one of those (L16) was found to induce a conformational change in reconstituted particles (Teraoka and Nierhaus, 1978), suggesting a structural role in subunit function.

In the early 1980s, the discovery of RNA molecules (ribozymes) with catalytic activity (the RNA moiety of the ribonucleoprotein RNaseP, and the self-splicing *Tetrahymena* 26S rRNA intervening sequence; reviewed in Cech and Bass, 1986) lent credence to the idea that large subunit (23S) rRNA might play a direct role in catalyzing peptide bond formation. Although some researchers had previously suspected the role of rRNA in ribosome function, the discovery of ribozymes revolutionized thinking about rRNA (Moore, 1988). The idea of an rRNA peptidyl transferase gained acceptance during the 1980s and 90s due to strong experimental support. Evidence included the retention of peptidyl transferase activity by *Thermus aquaticus* 50S subunits subjected to multiple rounds of proteinase K, detergent (SDS), and phenol extractions that removed ~95% of the subunit r-proteins (Noller et al., 1992); in contrast, treatment of the same *T. aquaticus* subunits with ribonuclease (RNase T1) easily abolished peptidyl transferase activity (Noller et al., 1992). A more extensive review of the evidence supporting the roles of rRNA in translation, including peptidyl transferase, was published by Noller

(1991), summarizing roughly two decades of research on functional interactions between ribosomal constituents and substrates in the prokaryote system.

1.2.7. 1980s-2000s: Crystallography and a structural marvel: the ribosome is a ribozyme

To gain a comprehensive understanding of the translation functions of the ribosome, researchers are determining the atomic structure of this enormous complex, once again focusing largely on the relatively small prokaryotic 70S ribosome. The structural biology-based renaissance in ribosome research is largely due to the advent of improved technologies (x-ray crystallographic techniques, software, synchrotron light sources) allowing the detailed mapping of subunit structure. The biggest hurdle in crystallography is growing good quality crystals; for a macromolecular complex as large as a ribosome or a ribosomal subunit, this task is immense. Attempts to crystallize the ribosome and its subunits date back to the late 1970s and early 1980s with work on *Bacillus stearothermophilus* (commentary in Pennisi, 1999; reviewed in Ramakrishnan and Moore, 2001). In the search for stable crystals, both a switch in organism and technique was required; the most stable crystals to date have been obtained with archaebacterial and thermophilic bacterial ribosomes (from the halophilic archaebacterium *Haloarcula marismortui* and the thermophilic bacterium *Thermus thermophilus*) that have been cryocooled to last longer under x-ray diffraction (Pennisi, 1999; Ramakrishnan and Moore, 2001). By 1999, the *H. marismortui* 50S subunit structure had been mapped to 5 Å (Ban et al., 1999), and the *T. thermophilus* 30S subunit (Clemons et al., 1999), and 70S ribosome (Cate et al., 1999), had been solved to 5.5 Å and 7.8 Å resolution, respectively.

In 2000, structural studies of the 70S ribosome came to an exciting fruition; the structures of both the *T. thermophilus* 30S subunit (Wimberly et al., 2000) and the *H. marismortui* 50S subunit (Ban et al., 2000) were resolved at 3 and 2.4 Å, respectively. This accomplishment included a demonstration (through binding of substrate analogs to the 50S subunit) that the peptidyl transferase function is a function of RNA, making the ribosome a ribozyme (Nissen et al., 2000). The 3 Å 30S subunit structure and its functional companion paper detailing interactions of three different antibiotics with the subunit (Carter et al., 2000) also shed light on ribosome function in terms of interaction

with mRNA and tRNA, the decoding process, and tRNA translocation. Researchers have continued to mine crystal structure data (e.g. Steitz, 2005) for information on ribosome and antibiotic function, as well as add to the bank of published structures related to ribosomal subunits (e.g. wobble base pair in the 30S subunit to 3.05 Å, Murphy and Ramakrishnan, 2004). Knowledge of subunit structure has also prompted experiments evaluating the functional roles of individual r-proteins (Hoang et al., 2004).

Elucidating the atomic structures of 80S ribosomes and their subunits remains a daunting challenge; at ≥4 MDa, eukaryotic ribosomes are almost twice the size of their ~2.4 MDa 70S counterparts; however, cryo-electron microscopy has proven useful for studying these large ribosomes in a variety of functional states (Ramakrishnan and Moore, 2001). In the mid- to late 1990s, mammalian (rabbit) 40S subunit (Srivastava et al., 1995), and whole wheat germ (Verschoor et al., 1996) and yeast (Verschoor et al., 1998) ribosome structures were resolved at ~55, 38 and 35 Å, respectively, and comparisons between the 80S ribosomes as well as their cryo-EM reconstructed *E. coli* 70S counterpart indicated several conserved structural features. While all three ribosomes (yeast, wheat germ, *E. coli*) share overall similarity in terms of primary morphological domains characterizing the subunits, the 80S ribosomes are broader and more elliptical than the 70S (although yeast ribosomes appear more compact and globular than those of wheat), and some morphological features, such as the 'spur' at the bottom of the 30S subunit, are not shared between 70S and 80S ribosomes (Verschoor et al., 1996; Verschoor et al., 1998). As with ribosomes from other eukaryotes (yeast, mammals, wheat germ), the recent cryo-EM structure and proteomic analysis of the *Chlamydomonas reinhardtii* 80S ribosome at 25 Å (Manuell et al., 2005) confirms the high level of conservation between the algal cytoplasmic ribosome and those from other eukaryotic species.

Cryo-EM reconstructions of ribosomes in functional states have been used successfully for both 70S (*E.coli*) and 80S (yeast, rabbit, canine, rat) ribosomes to examine the active processes of the ribosome such as elongation and cotranslational peptide translocation. It has been suggested that in active states of the ribosome, where ligands (e.g. substrates, substrate analogs, antibiotics) are present, greater structural stability can be achieved, leading to better data resolution (Gomez-Lorenzo et al., 2000).

Consequently, the highest resolution cryo-EM structures of *E. coli* (20 Å, Agrawal et al., 1998; 17.5-18.4 Å, Agrawal et al., 1999) and yeast (17. 5 Å, Gomez-Lorenzo et al., 2000) ribosomes have been achieved through the binding of elongation factors (EFs; *E. coli*: EF-G, yeast: EF-2) and stabilizing ligands (ribosome substrates and drugs; *E. coli*: GDP/GTP analog, tRNAs, fusidic acid; yeast: sordarin) in the study of elongation. Yeast 80S ribosomes engaged in co-translational peptide translocation and complexed with the translocon (peptide channel) of the ER have also been the subject of cryo-EM reconstruction at 25 Å (Ménétret et al., 2000) and 15.4 Å (Beckmann et al., 2001); rabbit and canine translocating ribosome complexes (canine translocons) have been visualized at 25-29 Å (Ménétret et al., 2000); the canine ribosome/translocon complex has also been resolved to 17 Å (Morgan et al., 2002). These cryo-EM studies have attempted to capture the dynamic nature of the ribosome; functional states too transient or difficult to crystallize for x-ray crystallography can be visualized using this technique.

1.2.8. 1980s-2000s: Biogenesis, genomics, and the return of ribosomal proteins

While the mechanism of translation has occupied most ribosome researchers for the past forty years, other questions remain critical to understanding the extraordinary complex at the center of this process. A cadre of new and improved tools including high-throughput sequencing, advanced molecular techniques, genomics, and proteomics, is now unraveling the ribosome, its constituents, its evolutionary history, and its connections to other processes within the cell. Beyond the study of translation, recent research has focused on aspects of ribosome biogenesis including rRNA and r-protein gene expression (e.g. Arabi et al., 2005; Grandori et al., 2005; Grewal et al., 2005; Perry, 2005), rRNA processing (e.g. reviewed in Granneman and Baserga, 2004), r-protein modification (e.g. Bloemink and Moore, 1999; Nusspaumer et al., 2000; Hanson et al., 2004), subunit formation (e.g. Saveanu et al., 2003; Schäfer et al., 2003; Stagg et al., 2003), and subunit transport (e.g. Trotta et al., 2003); knowledge of these processes has steadily increased since the 1980s, particularly in *E. coli* and yeast. While structural studies of the ribosome itself serve to reinforce the many similarities between ribosomes across kingdoms, studies of the regulation of rRNA and r-protein gene expression and ribosome biogenesis often highlight the divergence in regulatory mechanisms across

species, even when the functional sequences and structures of ribosomal constituents are highly conserved.

While bacterial ribosomes have dominated structural and translation studies, current ribosome biogenesis research primarily focuses on the budding yeast, *Saccharomyces cerevisiae*. Dating back to the late 1960s and early 1970s, studies of bacterial ribosome formation have primarily dealt with *in vitro* reconstitution experiments; functional subunits could be obtained using constituent rRNA and r-proteins (from *E. coli* or the thermophile *Bacillus stearothermophilus*) under reconstitution conditions, including a high incubation temperature (reviewed in Nomura, 1990). These initial *in vitro* reconstitution experiments were followed by numerous experiments where the step-wise addition of r-proteins onto rRNA identified primary, secondary, and tertiary binding r-proteins (reviewed in Nierhaus, 1991; Culver, 2003). The requirement for a heating step in the *in vitro* reconstitution experiments suggested the existence of other assembly factors *in vivo* (Culver, 2003; Williamson, 2003); to date, a handful of extraribosomal factors involved in bacterial *in vivo* subunit biogenesis (in addition to rRNA processing) have been identified, including RNA helicases such as CsdA (Cold-shock DEAD box protein A), involved in biogenesis of the 50S subunit (Charollais et al., 2004), and protein chaperones such as DnaK (reviewed in Culver, 2003; Williamson, 2003).

In contrast to the relative autonomy of prokaryotic ribosome synthesis, eukaryotic ribosome biogenesis, currently studied primarily in yeast, not only involves numerous non-ribosomal factors and multiple intermediate pre-subunit particles, but also takes place in multiple locations within the cell; subunit biogenesis is generally studied through the isolation of pre-subunit intermediate particles as they move from nucleolus, to nucleus, to cytoplasm. Studies on the assembly steps from 90S pre-ribosomes to 60S and 40S subunits in yeast have identified at least ~35 non-ribosomal proteins associated with 90S pre-ribosomal particles, ~50 non-ribosomal proteins associated with various pre-60S particles, and ~14 (8 major) non-ribosomal proteins associated with various pre-40S particles (Nissan et al., 2002; Schäfer et al., 2003). Many of these non-ribosomal factors involved in subunit synthesis have also been found associated with several steps in the cell cycle, suggesting connections between the two processes (Dez and Tollervey,

2004). In addition to the non-ribosomal proteins associated with nascent subunits, numerous ribonucleoprotein complexes are required for earlier steps in assembly such as RNA modification (reviewed in Lafontaine and Tollervey, 1998; Fromont-Racine et al., 2003). Recent research on the hundreds of non-ribosomal RNA and protein moieties involved in rRNA processing and subunit biogenesis has been reviewed by Fatica and Tollervey (2002) and Fromont-Racine and colleagues (2003).

Research on ribosomal constituents has been greatly aided by the sequencing of entire genomes over the past ~15 years. Not only has this simplified the isolation and cloning of rRNA and r-protein genes for downstream applications, but now genome-wide comparisons can be made identifying r-protein genes, via sequence homology, from species across all kingdoms. Instead of being restricted by 2D gel electrophoresis comparisons, made difficult by differences in technique (McConkey et al., 1979), evolutionary relationships and common ribosomal component features, including secondary structures, can be determined through sequence comparison and modeling software. Due to the highly conserved nature of rRNA and most r-proteins, sequence identity itself is a useful tool; for example, the sequenced r-protein set (79 r-proteins) from rat (*Rattus norvegicus*) is often used as a standard for comparison to identify r-proteins in other multicellular eukaryotes (e.g. Arabidopsis, Barakat et al., 2001; catfish, *Ictalurus punctatus*, Patterson et al., 2003; human, Kenmochi et al., 1998). Comparison of rRNA sequence and structure, and homologous r-proteins, from archaebacteria, eubacteria, and eukaryotes has revealed a remarkable conservation of core structural features across all major domains of life (Mears et al., 2002).

On the periphery since the discovery of ribozymes in the 1980s and the enzymatic functions of rRNA, the diverse array of r-proteins has begun to re-emerge as an area of research interest. R-proteins have resurfaced in two main ways: as structurally or functionally important to translation (e.g. S5, Ramakrishnan and White, 1992), or from mutant screens, cDNA libraries, and global gene expression profiling (e.g. microarrays) as being involved in diverse stress and developmental processes. R-protein genes have been identified in screens for genes upregulated in cancer cells (Kim, J.-H. et al., 2004), in response to cold treatment (Sáez-Vásquez et al., 2000; Kim, K.-Y. et al., 2004) or UV radiation (Casati and Walbot, 2003), during various stages of development

(Taylor et al., 1992; Van Lijsebettens et al., 1994), and numerous other conditions. In addition, several r-protein genes have been mapped to disease loci in humans (Uechi et al., 2001) and others have been shown to cause disorders such as the *Minute* phenotype, first reported in *Drosophila* (Kongsuwan et al., 1985), and subsequently in plants (Weijers et al., 2001). As a result, a new awareness has emerged concerning the unique origins of r-proteins and the possibilities of extraribosomal functions (Wool, 1996), or extra-translational functions of r-proteins within the ribosome (e.g. translocon docking and chaperone interaction, Beckmann et al., 2001; Kramer et al., 2002). This renewed interest in r-proteins has led to the creation of the Ribosomal Protein Gene (RPG) database for prokaryotic and eukaryotic r-protein sequence, structural, and taxonomic information (Nakao et al., 2004).

1.3. Basic Ribosome Structure

To summarize the discoveries reviewed above, ribosomes are ribonucleoprotein complexes comprised of 2 subunits (Figure 1.1), "large" (L) and "small" (S), which associate to form the complete ribosome on a molecule of mRNA; these may be present in the cytoplasm or associated with membranes of the ER. Studies in bacteria and cyanobacteria have determined that prokaryotic ribosomes have sedimentation coefficients of approximately 70 Svedberg units (S), as do plastid and mitochondrial ribosomes (Spirin and Gavrilova, 1969). The 70S ribosome is comprised of 30S and 50S subunits; the subunits of *E. coli* ribosomes are approximately 900 and 1600 kDa, combining to form the ~2.4-2.5 MDa prokaryotic ribosome (Boublik et al., 1990). Eukaryotic ribosomes (yeast, mammals, and plants) sediment at about 80S, with 40S and 60S subunits; the 80S ribosome is approximately 3-4.5 MDa, although actual molecular weights and sedimentation values can vary, depending upon species and tissue type examined (Bielka, 1982). In both prokaryotes and eukaryotes, the peptidyl transferase activity is localized to the large ribosomal subunit and mRNA binding, translocation, and tRNA anticodon/mRNA codon matching functions have been localized to the small subunit. It is clear from high resolution crystal structures that the size of the ribosome, calculated as ~10-20 nm from its appearance as a solid particle using electron microscopy, has been underestimated; the ribosome from the bacterium *T. thermophilus* measures ~210 Å (21 nm) across (Cate et al., 1999) and the *H. marismortui* large

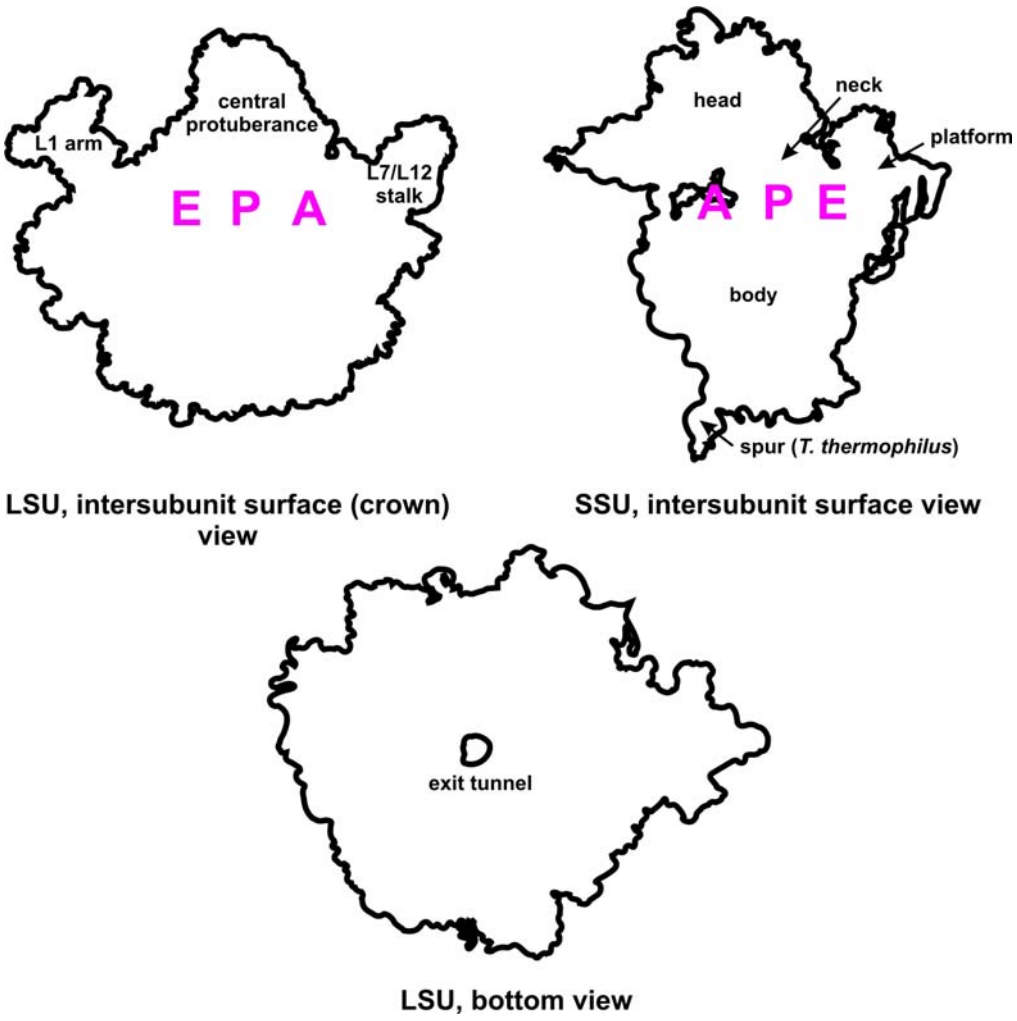


Figure 1.1. Basic ribosome morphology showing major structural features, based on the prokaryotic ribosome. Crown view of large subunit (LSU) and front view of small subunit (SSU) showing intersubunit interface. Approximate positions of A, P, and E sites indicated in pink lettering. Bottom view of LSU shows polypeptide exit tunnel. SSU diagram modified from Brodersen et al. (2002); spur at bottom is a feature of the *Thermus thermophilus* 30S subunit, not found in 80S ribosomes. LSU diagram modified from Klein et al. (2004).

subunit alone is ~250 Å across (Ban et al., 2000). Eukaryotic ribosomes, almost twice the mass of prokaryotic ribosomes, increase their dimensions via rRNA expansion segments and an increase in r-protein number (see below).

1.3.1. rRNA

Approximately 45-55% of each ribosomal subunit is rRNA (Bielka, 1982); it is rRNA that makes up the bulk and overall shape of each subunit (Ban et al., 2000; Wimberly et al., 2000). The small subunit has a single molecule of 16S-18S rRNA organized into four secondary structure domains (5' domain, central domain, 3' major domain, 3' minor domain; Figure 1.2) that make up the three major morphological domains of the small subunit (body, platform, head; Figure 1.1; Wimberly et al., 2000). The 16S rRNA is approximately 1500 nucleotides (nt) in length and the 18S rRNA of the 40S subunit is ~1800-2300 nt in length (Raué et al., 1988). The large subunit contains either 2 or 3 molecules of rRNA; the 50S subunit contains a 23S (Figure 1.3) and a 5S rRNA, while in eukaryotes, the 60S subunit contains a 25-26S (e.g. Arabidopsis, yeast) or 28S (e.g. rat, mouse) rRNA, and a molecule each of 5.8S and 5S rRNA. The 5.8S rRNA shows homology to the 5' end of the 23S rRNA (Nazar, 1980; Raué et al., 1988). Plastid (70S) ribosomes contain 23S, 16S, and 5S-type rRNAs; in the chloroplast ribosomes of vascular plants, there is a 4.5S rRNA corresponding to the 3' end of the 23S rRNA (Raué et al., 1988; Subramanian et al., 1990).

The rRNA of the large subunit has a modular nature: the 5.8S and 4.5S rRNA moieties represent portions of dissected 23S-like and 23S rRNAs, respectively; because of this modular nature, the 23S-like (eukaryotic) and 23S rRNAs along with 5.8S or 4.5S rRNAs are often referred to collectively as large subunit (LSU) rRNA (Raué et al., 1988; Schnare et al., 1996). LSU rRNA (including 5.8S rRNA where applicable) has six secondary structure domains (I-VI; Figure 1.3) and ranges from ~3000 nt in length in prokaryotes and the eukaryotic *Giardia* species to ~3500-5000 nt in most eukaryotes (Raué et al., 1988; Schnare et al., 1996). 5S rRNA can be considered a seventh structural domain in the large subunit (Steitz and Moore, 2003). The crystal structure of the 50S subunit shows that, unlike the rRNA of the small subunit, 23S and 5S rRNA secondary structures do not correlate well with individual morphological features of the large

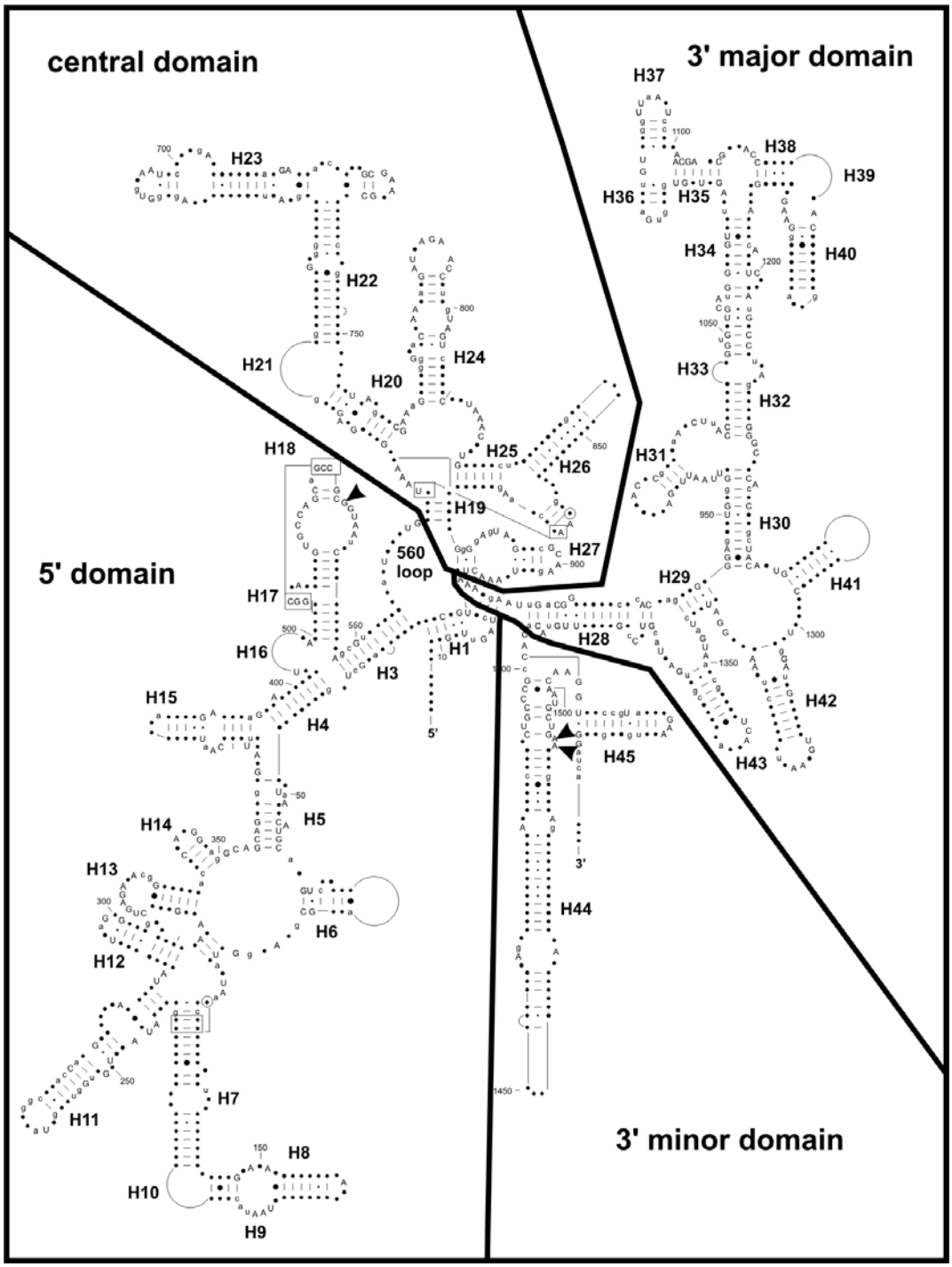


Figure 1.2. SSU rRNA, secondary structure and nucleotides conserved across archaea, bacteria, and eukarya, superimposed onto the *E. coli* structure, *E. coli* nucleotide numbering. Adapted from Mears et al. (2002) supplementary figure, Comparative RNA Web site (http://www.rna.icmb.utexas.edu; Cannone et al., 2002). Helix numbering as modified in Brodersen et al. (2002).

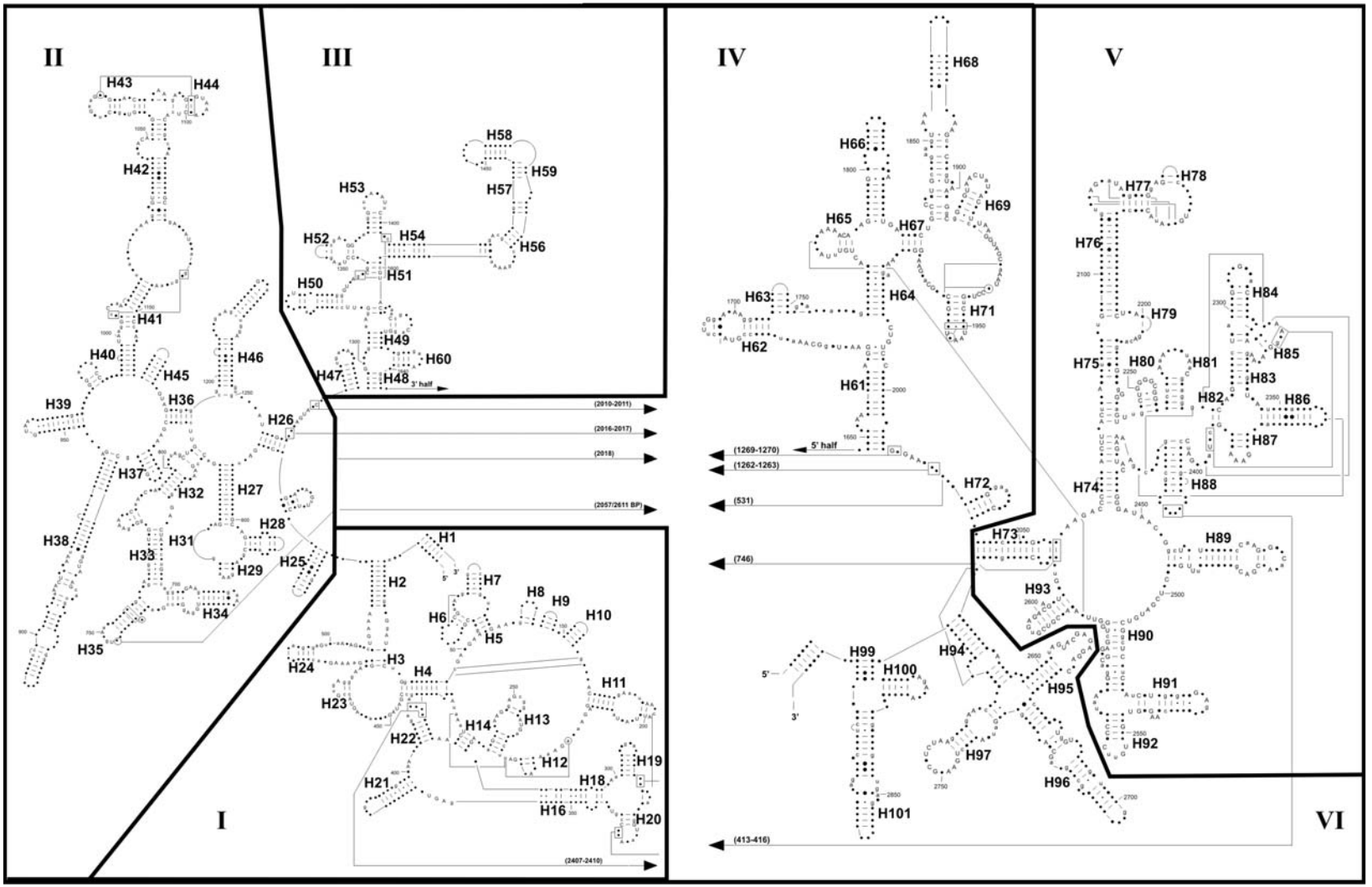


Figure 1.3. LSU rRNA, secondary structure and nucleotides conserved across archaea, bacteria, and eukarya, superimposed onto the *E. coli* structure, *E. coli* nucleotide numbering. Arrows indicate tertiary interactions between 5' and 3' halves of the rRNA. Adapted from Mears et al. (2002), Comparative RNA Web site (http://www.rna.icmb.utexas.edu; Cannone et al., 2002). Helix numbering as modified in Klein et al. (2004), domains I-VI divided as in Ban et al. (2000).

subunit; instead, the domains of LSU rRNA interact to form an enormous, singular, complex mass (Ban et al, 2000).

1.3.1.1. rRNA genes (rDNA)

Both prokaryotic and eukaryotic rRNA genes are arranged in clusters. *E. coli* has seven rDNA clusters consisting of 16S and 23S rRNA genes separated by two tRNA- and 5S rRNA-containing spacer regions, with each operon containing single-copy 16S and 23S genes and a variable number of tRNA and 5S rRNA genes; the arrangement of 16S and 23S rRNA genes as a single transcription unit with spacer tRNA genes is common among eubacteria (reviewed in Srivastava and Schlessinger, 1990). In eukaryotes, the rRNA genes are generally arranged in transcription units of $^{5'}$18S – 5.8S – 25/26/28S$^{3'}$, with the genes separated by internal transcribed spacers (ITS); these transcription units are arranged in tandem repeats, each preceded by a short external transcribed spacer (ETS), and separated by non-transcribed spacers (reviewed in Hadjiolov, 1985; Sollner-Webb and Tower, 1986). The number of repeat units was estimated in a variety of species in the 1970s and 80s using saturation hybridization of a labeled rRNA probe to filter-bound DNA; the values obtained range from hundreds (e.g. 100-200 in mammals, fungi, and some arthropods) to thousands (e.g. 2,000-5,000 in some amphibian species; 2,000-13,000 in plants) of repeats per haploid genome (reviewed in Long and Dawid, 1980; Hadjiolov, 1985). It is these tandem repeats, found in specific regions on one or few chromosomes, which form the core of nucleolar organizing regions (NORs). In Arabidopsis, for example, the two NORs are found near the ends of chromosomes 2 and 4 and are composed of ~350-400 10 kb rRNA gene clusters for a total of ~3.5-4 Mb (AGI, 2000).

The 5S rRNA genes are also found in tandem arrays in eukaryotes but, with few exceptions, and in contrast to the case in bacteria (see above), they are not located within the rRNA gene repeat regions. In Arabidopsis, the 5S rRNA gene arrays are located away from the 18S-5.8S-25S rRNA gene repeats, near the centromeres of chromosomes 3, 4, and 5 (AGI, 2000). 5S rRNA genes can range from hundreds (e.g. ~200 in *Neurospora crassa*, 100-200 in *Drosophila melanogaster*) to thousands (e.g. 9,000-24,000 in *Xenopus laevis*) of genes per haploid genome (reviewed in Long and Dawid, 1980). Interestingly the 5S rRNA genes in some eukaryotes, including the yeasts *S.*

cerevisiae and *S. carlbergensis* and the slime mold *Dictyostelium discoideum*, are located within the non-transcribed spacer regions preceding the 18S-5.8S-25S transcription units (reviewed in Hadjiolov, 1985; Warner, 1989).

1.3.1.2. rRNA sequence and *structure – phylogenetic comparisons*

As the functional core of the ribosome, rRNA is highly conserved across all domains of life: eubacteria, archaebacteria, and eukarya. Phylogenetic comparisons show that both large and small subunit rRNAs contain a core of highly conserved sequence and secondary structure interrupted at specific locations with variable regions (also called expansion segments) that differ in sequence and structure between species (Raué et al., 1988; Schnare et al., 1996; Mears et al., 2002). It is these variable regions that account for the major differences in rRNA lengths between species (e.g. eukaryotic LSU rRNA can range from ~2800 to over 5000 nt, Schnare et al., 1996) and can be visualized as extensions of 40S and 60S subunits in cryo-EM structure comparisons between eukaryotic and *E. coli* ribosomes (Dube et al., 1998a; Dube et al., 1998b; Gomez-Lorenzo, 2000) and between the ribosomes of different eukaryotic species (Morgan et al., 2000). The variable regions/expansion segments occur toward the outside surfaces of the subunits, away from functional centers involved in peptidyl transferase, decoding (matching mRNA codon to tRNA anticodon), major subunit interfaces, and peptide exit (Dube et al., 1998a; Dube et al., 1998b; Morgan et al., 2000). The functional importance of these variable regions is largely unknown; experiments deleting or extending expansion segments in the LSU rRNA of yeast and the ciliate *Tetrahymena thermophila* have determined that alteration of at least some expansion segments has no deleterious effect on growth or translation (reviewed in Schnare et al., 1996).

1.3.1.3. rRNA functions

Given the relatively short period in which rRNA has been considered more than just a scaffold for the assembly of r-proteins, it is ironic that more is now known about the functions of rRNA than of most r-proteins. rRNA plays a key role in every aspect of translation including mRNA initiator region selection, tRNA binding, proofreading, termination, and peptidyl transferase activity (reviewed in Noller, 1991; Moore and Steitz, 2002). As mentioned above, one of the biggest indicators of functional

importance for a given region of rRNA is its level of phylogenetic conservation. The regions involved in peptidyl transferase, mRNA binding, and interaction with translation factors have the highest level of universally conserved residues across all domains of life, including organellar ribosomes (Mears et al., 2002). To date, prokaryotic rRNA has been much better characterized in terms of function than eukaryotic rRNA and so it will be the primary focus of this section. Given the extent of conservation of rRNA functional regions, it is expected that eukaryotic rRNA functions in much the same way as its prokaryotic counterparts.

The catalysis of peptide bond formation is now known to be a function solely of LSU rRNA (Nissen et al., 2000), but the catalytic mechanism is still a matter of debate (for a detailed discussion of peptidyl transferase, see Appendix 1). The peptidyl transferase active site of the ribosome has been localized to the central loop of domain V of 23S rRNA; domain V nucleotides 3' and 5' to the central loop form the A (aminoacyl) and P (peptidyl) sites of the large subunit. Crystal structures of 50S subunits complexed with substrate analogs have confirmed that the 3' CCA end (amino acid-carrying end) of aminoacyl tRNA interacts with nucleotides of the LSU rRNA A loop in the A site, and the CCA end of peptidyl tRNA (carrying a nascent peptide) interacts with the P loop of LSU rRNA in the P site, positioning the substrates for catalysis (Nissen et al., 2000; Hansen et al., 2002). The acceptor stems of tRNA occupying the A and P (and E, exit) sites are located at the bottom of a cleft in the subunit interface surface of the large subunit; the CCA termini of the A- and P-site tRNA acceptor stems are positioned at the mouth of the peptide exit tunnel that leads through to the back of the large subunit (Nissen et al., 2000). Although the exact mechanism of peptidyl transferase is not yet clear, it is known that it is purely an RNA-based mechanism, involving substrate positioning in the A and P sites (Nissen et al., 2000; Hansen et al., 2002), pH-dependent conformation or acid/base catalysis (Maden and Monro, 1968; Nissen et al., 2000; Katunin et al., 2002; Moore and Steitz, 2002; Steitz and Moore, 2003), and substrate-assisted catalysis via the A76 2'OH of the peptidyl tRNA (Weinger et al., 2004; Steitz, 2005). A combination of these factors contributes to the estimated 10^5- (Rodnina and Wintermeyer, 2003) to 10^7-fold (Sievers et al., 2004) increase in reaction rate of peptide bond formation on the ribosome as compared to model uncatalyzed reactions (i.e. ester

aminolysis; e.g. dipeptide formation between amino acids and aminoacyl adenylates in solution).

In addition to the central function of peptidyl transferase, rRNA is also involved in decoding, the matching of mRNA codons with tRNA anticodons. Decoding is a function of the small subunit – the 30S A and P sites are referred to as the decoding center. While the amino-acid-carrying acceptor stem of tRNA interacts with the large subunit, the anticodon arm interacts with mRNA and the small subunit at the 30S A, P, and E sites. During selection of the translation initiation site on prokaryotic mRNA, the 5' Shine-Dalgarno ribosome-binding sequence of the mRNA base pairs with the 3' terminus of 16S rRNA (Shine and Dalgarno, 1974; Steitz and Jakes, 1975). This mRNA binding occurs near the E site, where the head and platform of the 30S subunit meet (Ogle et al., 2001); as a result, the mRNA 3' to the ribosome binding site is positioned across the small subunit, with its first codon in the P site. The 30S A, P, and E sites are located between the head and body of the subunit, and are each composed of rRNA from two or more secondary structure domains (Carter et al., 2000). As with the peptidyl transferase center, interactions between rRNA and RNA (mRNA and tRNA) substrates play an important functional role in the decoding center, discriminating between correct versus incorrect base pairing between codon and anticodon (Carter et al., 2000; Moore and Steitz, 2002). Binding of mRNA and cognate tRNA in the A site induces a conformational change in A site residues A1492, A1493, and G530 (which interact with different base positions in the codon-anticodon helix; *E. coli* numbering), allowing the decoding center to strictly monitor the first two bases of the codon-anticodon helix for Watson-Crick base pairing, while the third ("wobble") position is less stringent (Ogle et al., 2001).

Unlike the peptidyl transferase center, the decoding center also includes r-proteins; the C-terminal tails of r-proteins S13 and S9 contact the P site, and in the A site a loop of S12 lies close to the codon/anticodon helix and the C termini of S13 and S19 lie between the A and P sites, and near the anticodon stem loop, respectively (Carter et al., 2000). Despite the presence of r-proteins, the decoding center is still primarily rRNA and is predicted to function without a r-protein requirement (Moore and Steitz, 2002). Consistent with this prediction, *E. coli* mutants lacking the C-terminal tails of

both S13 and S9, presumably creating a protein-free P site, are viable, though slow-growing, suggesting that the r-proteins are dispensable for basic 30S P site function but required for optimal activity (Hoang et al., 2004). Unlike the A and P sites the E site, occupied by the deacylated tRNA that has exited the P site, is composed mostly of r-protein (Carter et al., 2000). While the P site, and to a slightly lesser extent, the A site, are highly conserved across all domains of life, the E site is quite poorly conserved (Mears et al., 2002).

The more ancient and basic the ribosomal functional center, e.g. the peptidyl transferase and decoding centers, the less r-protein tends to be involved in the site compared to the more recently evolved functional centers (e.g. E site; Moore and Steitz, 2002). The α-sarcin-ricin domain (SRD, Mears et al., 2002), identified as the binding site for the ribotoxins α-sarcin and the ricin A chain, is a functional center with a substantial protein component; however, the rRNA constituent of the SRD is universally conserved (Mears et al., 2002). This large subunit region is involved in the binding of elongation factors (prokaryotic EF-Tu, EF-G; eukaryotic EF2) during translation (reviewed in Wool et al., 2000). The SRD has also been implicated in translational fidelity (decoding), translocation (movement of peptidyl tRNA from A site to P site following peptide transfer), and translation rate (Wool et al., 2000; Panopoulos et al., 2004). Other known functional domains incorporate rRNA and varying amounts of r-proteins. Core intersubunit bridges formed between subunit interfaces are primarily (possibly exclusively) rRNA while peripheral bridges involve rRNA and r-proteins (Cate et al., 1999). The peptide exit tunnel is composed primarily of 23S rRNA but also contains contributions from three r-proteins (*H. marismortui* L4, L22, and L39e; Nissen et al., 2000), two of which (L4, L22) form a constriction in the exit tunnel and contribute (L22) to the regulation of elongation arrest in *E. coli* (Nakatogawa and Ito, 2002). Finally, surrounding the peptide exit tunnel at the back of the large subunit in *H. marismortui*, the membrane translocation channel docking site consists of rRNA and six r-proteins (L19, L22, L23, L24, L29, L31e; Ban et al., 2000; Nissen et al., 2000); prokaryotic and eukaryotic LSU rRNA (Prinz et al., 2000; Beckmann et al., 2001; Morgan et al., 2002) and eukaryotic r-proteins (Beckmann et al., 2001; Morgan et al., 2002) have also been implicated in ribosome-translocon connections.

1.3.2. Ribosomal proteins

A large array of proteins constitutes the remaining ~one third to one half of the mass of the ribosome. While highly conserved RNA-based mechanisms may be the core of ribosome function, many r-proteins also show a high degree of conservation between species and across the major domains of life: bacteria, archaea, and eukarya (Lecompte et al., 2002; Mears et al., 2002). As reviewed above (section 1.2.5, 1970s: two-dimensional electrophoresis and ribosomal proteins), the number of r-proteins in each ribosomal subunit for a given species was first based on 2D gel analysis; although ribosome composition must still be analyzed experimentally to confirm true ribosomal constituents, the sequencing of complete genomes has allowed the identification of putative r-proteins across species. Identified by 2D gel analyses (e.g. Gantt and Key, 1983; Sikorski et al., 1983) or sequence comparison and gene mapping (e.g. Barakat et al., 2001; Uechi et al., 2001), the number of ribosomal proteins in prokaryotic and eukaryotic ribosomes varies depending on the species. Protein content of the ribosome generally increases from bacteria to eukaryotes; a comparison of complete genomes across all three domains of life identified 57 different bacterial r-protein families, 68 archaeal r-proteins, and 80 eukaryotic r-protein types (Lecompte et al., 2002). For example, the *E. coli* ribosome contains 55 proteins, 21 in the small subunit, 34 in the large subunit (Kaltschmidt and Wittmann, 1970); the *S. cerevisiae* cytoplasmic ribosome has at least 78 r-proteins, 32 in the 40S subunit and 46 in the 60S subunit (Mager et al., 1997; Planta and Mager, 1998); mammalian ribosomes such as those of rat (Wool et al., 1995) and human (Uechi et al., 2001) have 80 r-proteins, 33 in the 40S subunit and 47 in the 60S subunit (Wool et al., 1995). Plant cytoplasmic ribosomes have been estimated to contain 75 to 92 r-proteins, depending on the species (reviewed in Bailey-Serres, 1998); it should be noted, however, that these numbers have been determined using numerous 2D gel techniques, which can produce a multitude of conflicting results (McConkey et al., 1979), and no standardized study has been carried out between plant species. In an alternative approach, a search of the entire Arabidopsis genome with known rat r-protein sequences has identified 80 putative Arabidopsis cytoplasmic r-proteins (Barakat et al., 2001).

R-proteins are generally small (10-20 kDa), basic, RNA-binding proteins ranging in mass from ~4 to 30 kDa in *E. coli* (Arnold and Reilly, 1999) and from ~3.4 to ~47 kDa in rat (Wool et al., 1995) and Arabidopsis (Barakat et al., 2001; Chang et al., 2005). With a few exceptions, the majority of r-proteins are basic, due to high lysine and arginine contents and a small proportion of aspartate and glutamate residues (Bielka, 1982; Wool et al., 1995; Barakat et al., 2001), resulting in a pI >8 (isoelectric point >8; Kaltschmidt and Wittmann, 1970; Wool et al., 1995; Barakat et al., 2001). The basic residues of r-proteins tend to be concentrated in internal, N-, or C-terminal extensions while any acidic residues tend to be present in globular domains that are positioned to the outside of the two subunit surfaces (Brodersen et al., 2002; Klein et al., 2004). From r-protein and subunit molecular weights, it has been determined that most r-proteins are present as single copies in the ribosome (Hardy, 1975; Bielka, 1982; Tal et al., 1990); crystal structures have supported this assertion (e.g. Ban et al., 2000; Wimberly et al., 2000). The small group of acidic r-proteins (pI ~4 to pI ~6; Wool et al., 1995; Barakat et al., 2001), which includes the prokaryotic L7/L12 and eukaryotic P proteins (P0, P1, P2, P3), are the only r-proteins present in multiple copies in the ribosome (discussed below).

1.3.2.1. R-protein conservation and phylogenetic comparisons

The idea that the ancestral ribosome was once composed entirely of RNA was, at one time, speculation (Crick, 1968) but is now the general consensus among ribosome researchers (Woese, 2001; Steitz and Moore, 2003). Given the abundance of data confirming the central importance of rRNA in ribosome function (see above, section 1.3.1.3, rRNA functions) it is probable that the ribosome emerged from the 'RNA World' with proteins added to the core structure over evolutionary time (Campbell, 1991; Moore, 1993; Poole et al., 1999; Steitz and Moore, 2003). While rRNA may be the most ancient molecule type in the ribosome, many r-proteins are also highly conserved. In a comparison of 66 complete genomes (45 bacterial, 14 archaeal, 7 eukaryotic) across all three domains of life, 32 r-protein families are universally conserved, 33 r-proteins are conserved between archaea and eukarya, and only 23, 1, and 11 r-proteins are specific to bacteria, archaea, and eukarya, respectively (Lecompte et al., 2002). As with rRNA secondary and tertiary structures, highly conserved r-proteins tend to have known involvement in ribosome assembly or functional centers such as

intersubunit bridges, the decoding center, or the peptide exit tunnel (Lecompte et al., 2002; Mears et al., 2002).

The conservation of not only sequence, but function, has been experimentally confirmed for a number of r-proteins via heterologous binding and complementation experiments. R-protein S14 (RPS14 in eukaryotes, RPS11 in prokaryotes) is universally conserved between phylogenetic domains (Lecompte et al., 2002; Mears et al., 2002) and has extensive rRNA contacts in the 30S subunit by means of its long N-terminal tail (Wimberly et al., 2000). Both human (Rhoads and Roufa, 1987) and *Drosophila* (Maki et al., 1990) RPS14 are able to complement a Chinese hamster ovary (CHO) RPS14 mutant cell line, rescuing a drug-resistant phenotype, and demonstrating functional equivalence of human, rodent, and insect RPS14 *in vivo*. Another universally conserved r-protein, bacterial RPL23 (yeast RPL25, RPL23a in other eukaryotes; Lecompte et al., 2002; Mears et al., 2002), has been identified as one of six r-proteins surrounding the peptide exit tunnel of the 50S subunit, bound to Domain III of LSU rRNA (Ban et al., 2000; Nissen et al., 2000). RPL23 and its homologues are known to act as docking sites for chaperones in both prokaryotes (Kramer et al., 2002; Gu et al., 2003) and eukaryotes (Pool et al., 2002), and are implicated in connections between the 60S subunit and the translocon of the ER (Beckmann et al., 2001). Consistent with the high level of conservation between RPL23 homologues, yeast RPL25 binds to the cognate binding site on *E. coli* 23S rRNA (El-Baradi et al., 1985) and mouse 28S rRNA (Jeeninga et al., 1996), while *E. coli* RPL23 (El-Baradi et al., 1987) and rat RPL23a (Jeeninga et al., 1996) bind to the homologous site on yeast 26S rRNA, *in vitro*. *In vivo*, a yeast *l25* mutant has been complemented independently with both rat RPL23a (Jeeninga et al., 1996) and Arabidopsis RPL23aA (McIntosh and Bonham-Smith, 2001; Chapter 2, this volume), demonstrating functional equivalence of the eukaryotic homologues. Yeast RPL28- (Mager et al., 1997; formerly L29) and RPL10- (Mager et al., 1997; formerly Qsr1p) deficient mutants have also been complemented by eukaryotic homologues, RPL28 with mouse RPL27a (Fleming et al., 1989; formerly L27'; new nomenclature according to Wool et al., 1995), and RPL10 with human and corn (*Zea mays*) RPL10 homologs (Dick and Trumpower, 1998; formerly QM; RPL10 standard nomenclature of Wool et al., 1995 and Mager et al., 1997). RPL28 is universally conserved among all

phylogenetic domains and RPL10 is conserved among eukaryotes and archaea (Lecompte et al., 2002).

1.3.2.2. R-protein structure and rRNA binding

Crystal structures of both subunits have allowed detailed observation of r-protein:rRNA and r-protein:r-protein interactions, characterizing specific r-protein contacts within the ribosome (Brodersen et al., 2002; Klein et al., 2004). High-resolution examinations of 20 of the 21 r-proteins of the *T. thermophilus* 30S subunit (S2-20 and the *Thermus*-specific protein Thx; excluding S1; Wimberly et al., 2000; Brodersen et al., 2002) and 27 of the 31 r-proteins of the *H. marismortui* 50S subunit (L2-6, L7Ae, L10e, L13-15, L15e, L18, L18e, L19e, L21e, L22-24, L24e, L29-30, L31e, L32e, L37Ae, L37e, L39e, L44e; excluding L1, L10-12; Ban et al., 2000; Klein et al., 2004) have shown that r-proteins are structurally diverse and differ in how they contact RNA and other r-proteins. Most r-proteins have at least one globular domain, and many have an extended internal loop, N-, or C-terminal tail; however, a few r-proteins have no globular domain and exist as extended domains interlaced between rRNA helices (Ban et al., 2000; Wimberly et al., 2000; Brodersen et al., 2002; Klein et al., 2004). The complex associations between r-proteins and multiple rRNA domains have necessitated the solution of complete subunit structures in order to determine accurate, contextual, r-protein structures. Extended r-protein domains are generally disordered when viewed outside their ribosomal context (e.g. see Ramakrishnan et al., 1995; Draper and Reynaldo, 1999; Mao et al., 1999, Tishchenko et al., 2001), and biochemical analyses (e.g. RNase protection assays, chemical crosslinking) cannot discern the numerous weak interactions, in addition to their major primary binding site, that r-proteins make within the subunit (Ban et al., 2000; Klein et al., 2004). In the cooperative environment of the ribosome, r-proteins both stabilize, and are stabilized by, their rRNA-binding interactions (Brodersen et al., 2002; Klein et al., 2004).

One of the most striking features of the ribosome is its asymmetry; not only are the subunits unequal in size, but the distribution of r-proteins over and within each subunit is asymmetrical, filling gaps and crevices in the rRNA bulk of each subunit while avoiding the functional rRNA cores. The 30S *T. thermophilus* r-proteins are predominantly found towards the back and periphery of the small subunit, away from

the intersubunit interface (Brodersen et al., 2002). A similar pattern is also found in the *H. marismortui* 50S subunit, with the r-proteins forming a lattice over the subunit surface, away from the intersubunit interface and peptidyl transferase sites (Klein et al., 2004). R-proteins with extended internal, N-, or C-terminal domains comprise 80% (16 of 20) and ~44% (12 of 27) of the characterized 30S (*T. thermophilus*; Wimberly et al., 2000; Brodersen et al., 2002) and 50S (*H. marismortui*; Ban et al., 2000; Klein et al., 2004) subunit r-proteins, respectively. These extended domains, rich in lysine and arginine residues, tend to be buried within the subunit, while the more rigid globular domains, which tend to have more acidic residues (aspartate, glutamate; but see below), are found at the surface of the subunit (Brodersen et al., 2002; Klein et al., 2004). The distribution of positively-charged residues in r-protein extended domains in both *T. thermophilus* and *H. marismortui* r-proteins reflects their close interaction with the negatively-charged phosphodiester backbone of rRNA (Allers and Shamoo, 2001; Brodersen et al., 2002; Klein et al., 2004). While the extended domains of *H. marismortui* 50S subunit r-proteins comprise only 18% of the total subunit protein, they cover 44% of the total rRNA surface area contacted by 50S r-proteins (Klein et al., 2004). As is characteristic of halophilic proteins (Elcock and McCammon, 1998), the globular domains of the r-proteins of *H. marismortui* are quite acidic (18.5% aspartate and glutamate content), compared to the extended domains of the same proteins (7% acidic residues; Klein et al., 2004). The resulting negatively charged surface of the *H. marismortui* ribosome is adapted for halophilicity, with the strong repulsive electrostatic interactions between acidic residues possibly preventing aggregation of subunits at high salt concentrations (Elcock and McCammon, 1998). A comparison between the 30S r-proteins of *T. thermophilus* and *E. coli* shows another species-specific adaptation: extended terminal tails of the thermophilic r-proteins, a probable adaptation for greater subunit stability (Brodersen et al., 2002). Despite a high level of sequence and structural conservation of r-proteins between species, species-specific adaptations are common.

While available crystal structures of r-proteins in their ribosomal context represent well-defined r-protein sets from only two highly specialized prokaryotes, these data have provided a wealth of information about RNA-binding strategies employed by r-proteins. Unlike DNA-binding proteins, which usually access specific base sequences

via the major groove of a B-DNA helix, RNA-binding proteins must recognize a variety of single-stranded and irregular secondary structures, often via the phosphodiester backbone (Ban et al., 2000; Allers and Shamoo, 2001; Steitz and Moore, 2003). An analysis of contacts made by six r-proteins (Protein Data Bank, PDB entries) showed that 60% and 18% of H-bond interactions were formed with phosphate oxygens (O-1/O-2) of the RNA backbone and ribose 2'OH groups, respectively; only 22% of H-bonds were base-specific contacts (Allers and Shamoo, 2001). Of the more than 560 contacts between the 20 characterized *T. thermophilus* 30S r-proteins and 513 16S rRNA nucleotides, over 370 (~67%) are non-base-specific and solely via the phosphodiester backbone (Brodersen et al., 2002). A number of non-base-specific interactions are also made between r-proteins and the 5S and 23S rRNAs of the *H. marismortui* 50S subunit; almost all of the 27 r-proteins characterized make major contacts through interactions solely with the phosphodiester backbone (Klein et al., 2004). In general, non-base-specific H-bonds between RNA-binding proteins, such as r-proteins, and ribose 2'OH groups are often non-residue-specific, involving the carbonyl or amide groups of the polypeptide backbone, or can involve lysine and arginine side chains (Allers and Shamoo, 2001). Lysine and arginine side chains can also be involved in H-bonding and long-distance electrostatic interactions with the O-1/O-2 atoms of the phosphodiester backbone (Allers and Shamoo, 2001). Electrostatic interactions between the lysine- and arginine-rich r-protein extensions and core rRNA elements are probably important contributors to subunit stability. To counter the high number of negative charges in rRNA phosphodiester backbones, numerous stabilizing forces (RNA:protein interactions; magnesium ion bridging between phosphates; RNA:RNA interactions such as base-pairing and the A-minor motif, where an A base from one helix is inserted into the minor groove of another) are required to fold the rRNA into compact subunits (Steitz and Moore, 2003).

Base-specific nucleotide contacts and recognition of secondary structure motifs are key elements by which r-protein:rRNA interactions gain specificity. Like other RNA-binding proteins, r-proteins can make base-specific H-bonds through the minor groove or a distorted (widened via the non-canonical base-pairing common to rRNA, or at helix termini) major groove of rRNA helices, at bulged nucleotides (which are

commonly recognized by a hydrophobic binding pocket), or between bases in hydrophobic crevices (Klein et al., 2004). In the *H. marismortui* 50S subunit, most frequent base recognition occurs via the minor groove of rRNA helices, where H-bonding contacts are most commonly formed between guanine N-2 atoms and carbonyl groups of the polypeptide backbone or acidic side chains (i.e. aspartate, glutamate; Klein et al., 2004). H-bonds between N-2 and carbonyl or asp/glu groups are relatively common among RNA-binding proteins that make G-specific contacts (Allers and Shamoo, 2001). Slender, elongated arginine- and lysine-rich r-protein extensions are the best suited domains for base contacts through the narrow major groove of rRNA helices (Klein et al., 2004).

R-protein recognition of specific rRNA secondary structures primarily occurs via electrostatic interactions with distortions of the phosphodiester backbone (Klein et al., 2004), but can involve numerous other interactions. One distinct secondary structure found in rRNA (as well as a few other characterized RNAs) is the kink-turn (K-turn), which is recognized by the kink in the phosphodiester backbone as well as through base-specific interactions. The K-turn is a helix-internal loop-helix motif with an ~120^{o} kink in the phosphodiester backbone that results in a turn in the helix; the two stems of the helix end up related to each other by a 120^{o} angle in a characteristic 3D structure recognized by specific RNA-binding proteins (Klein et al., 2001). The six K-turns of the *H. marismortui* LSU rRNA interact specifically with r-proteins L4, L7Ae, L10, L15e, L19e, L24, L29, L32e, and L37Ae, and the two K-turns of the *T. thermophilus* 30S subunit interact specifically with S11 and S17; the K-turns are all found at the periphery of the subunits, with one near the 50S peptide exit tunnel and one near a 30S subunit interface bridge, indicating possible functional roles for these structures (Klein et al., 2001). The rRNA K-turns are recognized by their binding r-proteins in a variety of ways but involve many of the recognition patterns described above (Klein et al., 2001).

While r-proteins may share some common principles of RNA binding, they vary widely in structure and binding strategy. The globular domains of r-proteins generally consist of α-helices and β-sheets with few distinct characteristics, although some have similar domain structures as other RNA-binding proteins. Klein and colleagues (2004) grouped the 50S r-proteins into six categories based on topology: antiparallel α+β

(βαββ), β-barrel, zinc-binding, α-helical, L15/L18e, and a group of r-proteins (mixed α+β r-proteins) with α helices and β sheets but no structure in common with any other r-proteins. Some of the 30S subunit r-proteins share a motif similar to the antiparallel α+β/βαββ 50S r-proteins, referred to as the α+β sandwich motif (Brodersen et al., 2002), which also show topological similarity to the four-stranded β-sheet of the RNA recognition motif (RRM or ribonucleoprotein/RNP domain) of the U1A spliceosome protein (Oubridge et al., 1994; Brodersen et al., 2002; Klein et al., 2004). Although some r-proteins share similar overall topology, they recognize rRNA in a variety of different ways (Brodersen et al., 2002; Klein et al., 2004). The 50S βαββ, zinc-binding, α-helical, and mixed α+β r-proteins all bind rRNA in different manners, and while the β-barrel r-proteins (L2, L3, L14, L21e, L24) are similar in that they tend to interact with rRNA via the loops connecting their β-strands, their overall topology, and modes of rRNA contact, vary with differences in polypeptide length (Klein et al., 2004). In contrast, evolutionarily related r-proteins may have similar modes of rRNA recognition but differ widely in topology (Harms et al., 2001).

The only r-proteins that interact in a similar fashion are those related by structural homology; in the 50S subunit, L15 and L18e share identity between C-terminal globular domains and make contact with highly similar types of binding sites (Klein et al., 2004). R-proteins that bind to K-turns show the same diversity as other r-proteins; only structurally homologous RNA-binding proteins (L7Ae, L30e, 15.5 kDa spliceosome protein) bind K-turns in a similar manner (Klein et al., 2001). S11 of the 30S subunit and L18 also interact in a similar way with K-turns that are part of their respective binding sites (Brodersen et al., 2002). Clearly, while r-proteins may share some structural similarities, they tend to reflect co-evolution with their binding sites (e.g. L23/L25; Metzenberg et al., 1993; Chenuil et al., 1997) rather than with other r-proteins.

1.3.2.3. R-protein ribosomal functions

1.3.2.3.1. Structure and assembly

The roles of r-proteins in various aspects of ribosome function is less obvious now than would have been predicted forty, or even twenty, years ago when the role of rRNA as the functional core of the ribosome was unknown. Given the key roles of rRNA in translation, most r-proteins have been presumed to serve predominantly

structural roles. The structural data gathered in the past five years indicate that the majority of r-proteins serve, at least in part, as folding and stabilizing molecules for the complex folds assumed by rRNA (Ban et al., 2000; Wimberly et al., 2000; Brodersen et al., 2002; Klein et al., 2004). The RNA-binding interactions of most r-proteins in both subunits bury a substantial portion of individual r-protein surface area; between 7.7% (S2) and 54.4% (Thx) of 30S r-protein surface area (Brodersen et al., 2002) and from 11.4% (L7Ae) to 65.5% (L37e) of 50S r-protein surface area (Klein et al., 2004) is buried in rRNA contacts. The requirement for stabilizing forces in the ribosome is clear: as discussed above, the rRNA of the ribosome, at ~1.5 to 5 kb (Raué et al., 1988; Schnare et al., 1996), must fold into compact subunits, forcing together the 1500-5000 negative charges of the phosphodiester backbones in complex tertiary interactions. R-protein:rRNA interactions (in addition to rRNA:rRNA interactions and rRNA:Mg^{2+} interactions), especially those involving charged r-protein extensions (see above), help counter these repulsive forces and stabilize tertiary interactions (Steitz and Moore, 2003). The number of contacts made between each r-protein and rRNA far exceeds the number of interactions made between r-proteins. On average, each protein of the *T. thermophilus* 30S subunit contacts 5.7 rRNA helices (2-10 contacts from 1-3 domains) but only makes 1.2 contacts (range of 0-3) with other r-proteins (Wimberly et al., 2000). *H. marismortui* 50S r-proteins contact an average of 2.5 rRNA domains (1-6 domains each with the exception of L12, which makes no contacts with rRNA) but only 1.2 (range of 0-3) other r-proteins (Ban et al., 2000).

As well as stabilizing rRNA structure in the assembled subunit, some r-proteins are known to play specific roles in subunit assembly, binding and stabilizing multiple rRNA secondary structures, forming the compact rRNA tertiary structures of the subunit and creating binding sites for subsequent r-proteins (Brodersen et al., 2002; Klein et al., 2004). Assembly maps detailing the sequential association of r-proteins with rRNA in each of the *E. coli* ribosomal subunits were initially established in the 1970s-80s. The first r-proteins to bind to rRNA independently of other r-proteins were labeled primary binders or assembly initiator proteins, followed by the secondary binders that require association of one or more primary binders to associate with the subunit; these are followed by tertiary binding r-proteins (Held et al., 1974; Nowotny and Nierhaus, 1982;

Röhl and Nierhaus, 1982; Herold and Nierhaus, 1987; Mandiyan et al., 1991). Current structural data has enhanced many of the original assembly map results. In the 30S subunit, the primary binding proteins S4, S7, S8, S15, S17, and S20 (Held et al., 1974; Mandiyan et al., 1991) are predominantly globular r-proteins that bind to multiple helices, often at multi-helix junctions, organizing the rRNA domains of the head and body and establishing inter-domain contacts (Brodersen et al., 2002). S7 interacts with several helices of the 3' major domain, S4 brings together helix 21 of the central domain with multiple helices of the 5' domain, S8 and S15 bind and organize multiple helices of the central domain (also validated by hydroxyl radical probing; Jagannathan and Culver, 2003), S17 interacts with distal regions of the 5' and central domains, and S20 binds helices of the 5' and 3' minor domains (Brodersen et al., 2002). Other r-proteins with extended domains must assemble on the rRNA before the subunit is fully formed (it is unlikely such long, charged extensions could be inserted into fully compact domains), and many of them aid in the bridging of different rRNA domains. The N-terminal extension of the secondary binder S12 connects 5' and central domains as it weaves from the interface side of the subunit to contact S8 and S17 at the back of the body of the subunit (Brodersen et al., 2002). S9, S10, and S14 have extensions that bring together the helices of the 16S 3' major domain in the head, and S2, S3, and S5 are required for joining the head of the subunit to the body (Brodersen et al., 2002). In contrast to the predominantly globular primary binding proteins of the small subunit, the prokaryotic large subunit assembly initiator r-proteins L24 and L3 (Nowotny and Nierhaus, 1982) and the early assembling r-proteins L2, L4, and L22 (Röhl and Nierhaus, 1982; Herold and Nierhaus, 1987) all have extended domains (Ban et al., 2000; Klein et al., 2004). L24 and L3 influence the binding of subsequent r-proteins via their effects on rRNA structure, not through direct contact with other r-proteins; L24 binds domain I exclusively, and is required for the formation of the tertiary structure of the domain, which influences the binding of the early assembly protein L22, and subsequently the binding of numerous other r-proteins due to its contact with all six 23S rRNA domains (Klein et al., 2004). The extensions of L3 contact every 23S domain except domain I, separating its assembly initiator role from that of L24 (Klein et al., 2004).

1.3.2.3.2. Central protuberance and the L1 stalk

While the role of r-proteins in subunit stability may give the impression of r-proteins as static entities, r-proteins are, in fact, a major part of the dynamic nature of ribosomes. The central and lateral protuberances of the large subunit are flexible structural features that play important functional roles in translation. The central protuberance is formed by 5S rRNA at the top of the large subunit. Other than a few backbone interactions between domain II (helices 38-40) of 23S rRNA and the helix 4/5 arm (including loops E and D) of 5S rRNA (Ban et al., 2000; Klein et al., 2004), 5S rRNA associates with 23S rRNA primarily through r-protein-mediated interactions; L5 and L18 form the primary contacts between the 50S subunit and 5S rRNA, with limited contributions from L10e, L21e, and L30 (Ban et al., 2000; Klein et al., 2004). Interactions between the central protuberance r-proteins and their rRNA binding sites primarily occur via the minor grooves of rRNA helices, which are usually recognized by a β-sheet of the r-protein (Klein et al., 2004). L5 and L18 share substantial protein:protein interactions as do L18 and L21e (Ban et al., 2000; Klein et al., 2004). Interestingly, in both *Bacillus stearothermophilus* and *E. coli*, L18, which interacts with rRNA via base-specific and backbone 2'OH H-bonds and electrostatic interactions, requires phosphorylation of a serine residue (Ser), and the presence of Mg^{2+}, at physiological pH for 5S rRNA binding; suggesting a possible regulatory role for L18 in ribosome assembly or another aspect of translation (Bloemink and Moore, 1999).

The L1 lateral protuberance (or L1 stalk), as its name implies, is primarily formed by the r-protein itself. Modeling this ribosomal feature at high resolution within its ribosomal context, however, has proven difficult due to the flexibility of the region (Cate et al., 1999; Ban et al., 2000; Harms et al., 2001; Nikulin et al., 2003; Klein et al., 2004). The structure of *Sulfolobus acidocaldarius* L1, bound to its cognate binding site from *T. thermophilus* rRNA, was recently resolved at 2.65 Å (Nikulin et al., 2003). L1 is a globular two-domain protein that interacts with a highly conserved binding site (Nevskaya et al., 2005) in domain V (helices 77 and 78) of 23S rRNA (and homologous sites in eukarya) but not with other r-proteins (Ban et al., 2000; Nikulin et al., 2003). The β-strands and loops of L1 interact with rRNA primarily through the phosphodiester backbone, forming at least five highly conserved H-bonds (Nikulin et al., 2003;

Nevskaya et al., 2005). Interestingly, there is no significant conformational difference in L1 in its RNA-bound and unbound states (Nevskaya et al., 2005). Not only is the L1 arm a conserved structural feature, it is also involved in the removal of deacylated tRNAs from the ribosome. Comparison of the vacant *D. radiodurans* 50S subunit to the *T. thermophilus* 70S ribosome with occupied A, P, and E sites indicated an ~30 Å difference between the positions of the L1 protuberance in the two structures, suggesting a possible role for L1 movement during release of the E-site tRNA (Harms et al., 2001). The role of the L1 stalk in translocation was confirmed by subsequent cryo-EM studies of the *E. coli* 70S ribosome showing a pivoting of the stalk by ~20 Å toward a deacylated P-site tRNA upon binding of elongation factor G (EF-G) + GTP; the data suggest a model where, following GTP hydrolysis (to render EF-G + GDP), the ribosome resumes an "open" conformation and the L1 stalk swings away from the P-site, pulling the deacylated tRNA towards the E-site and subsequent exit from the ribosome (Valle et al., 2003). L1 has also been crosslinked to L33 on *E. coli* 50S subunits and 70S ribosomes (Kirillov et al., 2002).

1.3.2.3.3. Acidic r-proteins and the L7/L12 stalk

The other major structural feature of the ribosome formed by r-proteins is the L7/L12 or acidic stalk – a lateral protuberance formed by acidic r-proteins in both prokaryotic and eukaryotic ribosomes. The acidic r-proteins are unique due to their amino acid composition, their presence as dimers in the ribosome, and the fact that most do not directly contact rRNA but instead interact with other r-proteins (reviewed in Wahl and Möller, 2002). The stalk complex consists of a base r-protein that binds rRNA (L10 in prokaryotes or P0 in eukaryotes) and four r-proteins (Subramanian, 1975), two dimers of L7/L12 in prokaryotes (L7 and L12 are identical except for the N-terminal acetylation found in L7; both forms can be called L12) or a tetramer of P1 and P2 proteins in eukaryotes (reviewed in Wahl and Möller, 2002; Hanson et al., 2004), that interact with the base r-protein. The L10 and L12 r-protein families are present in all phylogenetic kingdoms (Lecompte et al., 2002) and L12 r-proteins show a similar basic structure, with N- and C-terminal domains separated by a flexible hinge region (Köpke et al., 1992; Wahl and Möller, 2002). The L10/P0 proteins show conservation between species, and P0 proteins from *Aspergillus fumigatus*, *D. discoideum*, rat, and human will

all complement a *S. cerevisiae* P0-deficient strain (Rodríguez-Gabriel et al., 2000); human P0 incorporation into yeast ribosomes will even confer resistance to the anti-fungal agent sordarin (Santos et al., 2004). There are some significant differences between L10/P0 proteins, however, and while the (N-terminal) rRNA-binding domain of P0 proteins is conserved, they do not bind all heterologous P1/P2 proteins (Rodríguez-Gabriel et al., 2000). Accordingly, while the entire rat P0-P1/P2 stalk complex will bind to *E. coli* ribosomes stripped of their native L10-L7/L12 complex, the hybrid ribosomes will only bind eukaryotic EF-2, not prokaryotic EF-G (Uchiumi et al., 1999).

Sequence identity between the L12-like r-proteins is limited and even eukaryotic P proteins show structural and functional diversity between species, and between and within P protein families in single species (Mager et al., 1997; Szick et al., 1998; Zurdo et al., 2000; Barakat et al., 2001; Hanson et al., 2004). The four P proteins of *S. cerevisiae*, P1A, P1B, P2A, and P2B (A, B also known as α, β forms), show sequence (Mager et al., 1997), structural (Zurdo et al., 2000), and functional (Hanson et al., 2004) diversity, binding to P0 as two heterodimers (P1A/P2B, P1B/P2A). Plant acidic r-proteins have been identified in numerous species, including rice (Hihara et al., 1994), Arabidopsis (García-Hernández et al., 1994; Barakat et al., 2001), and maize (Szick et al., 1998), with the five Arabidopsis P2 r-proteins varying up to 14.1% between isoforms (Barakat et al., 2001). In addition, plants possess a unique P protein, P3 (Szick et al., 1998; Barakat et al., 2001), not found in other eukaryotes (Lecompte et al., 2002). Unlike bacterial acidic r-proteins, eukaryotic acidic r-proteins are found both associated with ribosomal subunits and in free pools in the cytoplasm (Köpke et al., 1992; García-Hernández et al., 1994; Wahl and Möller, 2002). The distribution of acidic r-proteins between these two states has been correlated with growth stage, with younger cells having less acidic r-protein in the cytoplasmic pool (García-Hernández et al., 1994).

The acidic stalk complex is another dynamic, flexible protuberance that has not been amenable to structure determination via high-resolution crystallography (Cate et al., 1999; Ban et al., 2000) or current cryo-EM models (Gao et al., 2003). The acidic stalk is part of the translation factor binding center, a relatively protein-rich portion of the ribosome which, in addition to the L10-L12 proteins, includes the conserved r-proteins L3, L6, L11, L13, and L14 (and the sarcin-ricin domain rRNA; Ban et al.,

2000; Klein et al., 2004). The stalk complex is involved in factor-binding and the stimulation of GTP hydrolysis by elongation factors (e.g. Uchiumi et al., 1999); L12 has been implicated in interactions with initiation, elongation, and release factors, as well as the stimulation of EF-G GTPase activity (reviewed in Wahl and Möller, 2002). The involvement of P1/P2 proteins in the regulation of translation selectivity has also been demonstrated with P1/P2 deficient *S. cerevisiae* strains showing different protein expression patterns compared to wild type strains (Remacha et al., 1995).

1.3.2.3.4. Decoding, translocation, and intersubunit contact

In addition to their roles in factor binding and P- to E-site tRNA translocation, r-proteins have also been implicated in decoding, intersubunit contact, and co-translational peptide folding and translocation. In the decoding center, as mentioned above, a portion of S12 lies close to the A site, near the codon/anticodon helix (Carter et al., 2000). A highly conserved loop of S12 forms a base-specific H-bond with 16S rRNA residue A1492 (*E. coli* numbering), and a Mg^{2+}-mediated interaction occurs between the 2'OH group of a 'wobble' position base and a S12 backbone carbonyl group (Ogle et al., 2001). The C termini of S13 and S9 are part of the P site (Carter et al., 2000). While *E. coli* S13 and S9 deletion mutants are viable, they do show a slow-growth phenotype, suggesting that optimal ribosome activity requires these r-proteins at the P site (Hoang et al., 2004). Although contact between S9 (eukaryotic S16) and P site rRNA is conserved in eukaryotes, S13 (eukaryotic S18) contact has only been observed in *T. thermophilus* (Spahn et al., 2001). On the 60S subunit, *S. cerevisiae* r-proteins L10e (eukaryotes and archaebacteria only) and L11 (prokaryotic L5) show contacts with the T loop and the TΨC stem of P-site tRNA, respectively (Spahn et al., 2001). Prokaryotic r-proteins S7 and S11 interact with each other and bind the minor groove of the E site tRNA anticodon stem-loop, possibly facilitating the dissociation of the tRNA from the ribosome (Carter et al., 2000). Disruption of the S7-S11 interaction in *E. coli*, a highly conserved interaction in bacteria, results in a decrease in translational fidelity by increasing frame shifting, stop codon readthrough, and codon misreading (Robert and Brakier-Gingras, 2003). Impaired ribosome function, resulting from the disruption of the S7-S11 interaction, is consistent with the involvement of S7 and S11 in the structural dynamics of the 30S subunit during translocation (Robert and Brakier-Gingras, 2003).

Related to ribosomal dynamics, several r-proteins contribute to intersubunit bridges, some of which must be adjusted and re-formed during translocation, and all of which must be formed, broken, and re-formed during subunit association and dissociation. The number and nomenclature of the intersubunit bridges is not entirely standardized; the nomenclature developed for the ten bridges observed by Cate and colleagues (1999) in the 7.8 Å resolution *T. thermophilus* 70S structure was expanded for thirteen bridge interactions characterized in the 11.5 Å resolution *E. coli* 70S ribosome (Gabashvili et al., 2000). More recently, twelve intersubunit bridges have been characterized in the crystal structure of the 70S ribosome at 5.5 Å resolution; of these 12 bridges, five (B2a, B2b, B2c, B3, B7a) involve solely rRNA:rRNA contacts, six (B1a, B4, B5, B6, B7b, B8) involve rRNA:r-protein contacts, and one (B1b) involves a r-protein:r-protein contact (Yusupov et al., 2001). Interestingly, 24 Å resolution structures of yeast and rabbit 80S ribosomes revealed that most of the intersubunit bridges are evolutionarily conserved, although a unique eukaryotic bridge connects a point adjacent to the L1 arm to the 40S platform, a feature expanded in eukaryotes largely via a rRNA expansion segment (Morgan et al., 2000). A more recent study of the yeast ribosome at 15.4 Å resolution identified four eukaryote-specific bridges (eB8-eB11), all of which involve small and large subunit proteins (Spahn et al., 2001).

Although bridges involving protein interactions tend to be more peripheral than the core rRNA-based bridges (Cate et al., 1999; Yusupov et al., 2001), the r-proteins involved in subunit contacts still play key roles. A comparison of 70S ribosomes in multiple functional states showed that translocation of tRNAs (A → P → E site) during translation involves a 'ratchet-like' motion between subunits. Upon binding of EF-G + GTP, the 30S subunit rotates counterclockwise about 6^{o} on the large subunit, breaking contact with the A-site tRNA anticodon, then returns to its starting position upon GTP hydrolysis (EF-G + GDP), during which time translocation occurs and the A site is vacated (Frank and Agrawal, 2000). This ratchet motion involves the movement of three bridges (B1a, b, and c, Gabashvili et al., 2000 nomenclature) between the head of the 30S subunit and the 50S central protuberance; bridge B1b involves contacts between L5 and S13 (grouped as a single bridge by Yusupov et al., 2001) and B1a is a bridge between 50S helix 38 (H38) and S13 (Valle et al., 2003). Upon EF-G binding, L5 binds

to different positions on S13 in bridges B1b and c, and H38 breaks contact with S13 and interacts with S19 (Valle et al., 2003). The role of L5 and S13 is particularly important given their contacts with P-site tRNAs (Valle et al., 2003; see above). Deletion of S13 in *E. coli* severely impairs growth, subunit association, initiation complex formation, and increases factor-independent translocation (Cukras and Green, 2005). Bridge B1b appears to be a critical intersubunit contact; N-terminal mutations of S13 that perturb B1b result in growth retardation and an increase in stop codon readthrough and frameshifting, whereas changes to B1a have less impact on ribosome function (Cukras and Green, 2005).

Another bridge r-protein, L2, essential for subunit association and involved in the association of A- and P-site tRNAs with the 50S subunit (Diedrich et al., 2000), has long been implicated in peptidyl transfer due to the decrease or abolition of this activity in subunits with a mutated L2 or lacking L2, respectively (Cooperman et al., 1995; Ühlein et al., 1998; Diedrich et al., 2000). Although L2 cannot play a direct role in peptidyl transfer, given that its closest side chain is 23.5 Å away from the 50S active site, it does have rRNA contacts near the central loop of domain V, and most likely plays a role in stabilizing the structure of the active site (Nissen et al., 2000). In addition to its role near the active site, L2 is part of bridge B7b (Yusupov et al., 2001; Gao et al., 2003), another dynamic contact broken in the EF-G-bound ribosome as L2 undergoes substantial movement (Gao et al., 2003). A cryo-EM examination of wheat germ ribosomes at 9.5 Å resolution showed that r-protein L30e contributes to bridges B4 (where it interacts extensively with S13) and eB9; comparisons of ribosomal active states showed that L30e bridge contacts are lost in eEF-2-bound ribosomes, demonstrating the dynamic nature of B4 and eB9 (Halic et al., 2005).

1.3.2.3.5. The exit tunnel and cotranslational processes

At ~100 Å in length in the 50S subunit (Nissen et al., 2000) the peptide exit tunnel can protect up to ~40 residues of an elongating peptide (Matlack and Walter, 1995), but with an average diameter of 15 Å, the tunnel imposes a barrier to most peptide folding, barely accomodating α-helices (Nissen et al., 2000). This folding, however, is a key component in recognition, by the ribosome, of transmembrane sequences (TMS; 20-24 nonpolar residues that are inserted into a lipid bilayer) present in

membrane-associated proteins (Johnson, 2004). A ribosome docked at the translocon of the ER can identify whether a nascent peptide is a membrane protein or a secretory protein based on whether or not the nascent peptide has a TMS (Liao et al., 1997). Once a TMS is identified by the exit tunnel, the ribosome triggers structural changes: the translocon pore is sealed off and the ribosome-membrane junction is opened, allowing exposure of the cytoplasmic domain of the nascent peptide to the cytosol during co-translational insertion into the membrane (Liao et al., 1997). A nascent chain without a TMS is fully extended within the exit channel, contacting only r-protein L4 (~40 kDa) at the narrowest point of the tunnel (~10 Å); a TMS, however, is folded into an α-helix by the ribosome (free or membrane-bound) inside the exit tunnel, and crosslinks to two additional r-proteins, L17 (~18 kDa; prokaryotic L22) and L39 (~7 kDa), when the translocon pore is sealed and the ribosome-membrane junction is opened (Woolhead et al., 2004; Johnson, 2004). Conformational rearrangements of the *D. radiodurans* peptide exit tunnel, notably the 'swinging out' of the conserved β-hairpin of L22 to expose nonpolar residues to the tunnel and creating a 'swinging gate,' upon binding of the macrolide antibiotic troleandomycin (TAO) suggests a possible mechanism for restriction of peptide movement through the tunnel (Berisio et al., 2003). Given the changes in L22 elicited by macrolide binding (Berisio et al., 2003), the possibility of a TMS eliciting similar conformational changes would explain the control of the ribosome over co-translational membrane protein insertion (Johnson, 2004). Contact between a TMS and L39 appears to result in the opening of the ribosome-translocon junction (Johnson, 2004).

The r-proteins surrounding the exit tunnel at the back of the large subunit also play a critical role in regulating the fate of nascent peptides. Six r-proteins, L19e, L22 (L17 in eukaryotes), L23 (L23a), L24 (L26), L29 (L35), and L31e, were initially identified as surrounding the exit tunnel (Ban et al., 2000; Nissen et al., 2000), with L39e subsequently added to this group (Klein et al., 2004). There is a molecule of the prokaryotic chaperone trigger factor (TF) docked on virtually every ribosome in an *E. coli* cell (Lill et al., 1988), and as such, it is the first r-protein to interact with nascent peptides (Buskiewicz et al., 2004; Maier et al., 2005). TF forms a "cradle," with a hydrophobic inner surface over the peptide exit tunnel to shield nascent peptides from

premature folding or aggregation but also leaving enough of an open shape to allow access by other chaperones (Ferbitz et al., 2004; Maier at al., 2005). The binding site for TF on both translating and empty ribosomes is adjacent to r-proteins L23 and L29 near the peptide tunnel exit (Blaha et al., 2003), and crosslinking studies have confirmed that L23 and L29 interact with TF, with L23 being essential for TF binding to the ribosome (Kramer et al., 2002). The recent crystal structure of TF confirms the previous studies: TF binds L23, L29, and domain III of 23S rRNA via a single N-terminal 'tail' (Ferbitz et al., 2004; Maier at al., 2005).

The ribonucleoprotein signal recognition particle (SRP, Ffh in prokaryotes) which, in concert with its receptor (SR, FtsY in prokaryotes), guides nascent secretory and membrane proteins to the plasma membrane (prokaryotes) or ER (eukaryotes) for translocation (reviewed in Stroud and Walter, 1999; Keenan et al., 2001) is another chaperone that interacts with r-proteins. In eukaryotes, elongation arrest, which occurs following SRP/nascent chain binding (Keenan et al., 2001), is the result of competition at the factor binding center of the acidic stalk base of the large subunit, between the SRP Alu domain and elongation factors (Halic et al., 2004). Crosslinking studies in prokaryotes have shown that Ffh interacts with L23, which also interacts with TF, adjacent to the peptide exit tunnel (Gu et al., 2003; Kramer et al., 2002). In eukaryotes, crosslinking has also been observed between SRP and L23a and L35 (L23 and L29 in prokaryotes; Pool et al., 2002). The crystal structure of eukaryotic SRP docked to the ribosome confirms that SRP54, a subunit of SRP, contacts L23a and L35 (Halic et al., 2004). Given that TF, SRP, and membrane channels (see below) all dock via interactions with L23 and L29 (eukaryotic L23a and L35), it appears that these r-proteins form a general interaction site for numerous competing factors involved in cotranslational processes (Halic et al., 2004; Maier et al., 2005). Although, in prokaryotes, the Ffh/L23 interaction is unaffected by the presence of FtsY (Gu et al., 2003), in eukaryotes the binding of SRP54 to L23a is abolished, and binding to L35 enhanced, in the presence of SR and GTP, suggesting that a repositioning of SRP54 occurs as SRP54 and SR interact and bind GTP (Pool et al., 2002). In prokaryotes, Ffh and the chaperone TF can bind simultaneously to the ribosome and TF is only displaced when FtsY joins the complex (Buskiewicz et al., 2004).

Cryo-EM investigations of the ribosome/translocon interaction in yeast and mammals show the ribosome to position the peptide exit tunnel directly over top of the translocon pore (Beckmann et al., 1997; Ménétret et al., 2000; Beckmann et al., 2001; Morgan et al., 2002). There are four connections between the ribosome and the translocon, probably involving r-proteins L19, L23a, L26, and L35 (L19e, L23, L24, and L29 in prokaryotes; Beckmann et al., 2001; Morgan et al., 2002). These ribosome-translocon connections remain stable following peptide release and only dissociate upon translation of signal-less peptides (Potter and Nicchitta, 2000, 2002).

The SRP-nascent chain-ribosome complex docks at the plasma membrane or translocon via a GTP-dependent SRP:SR interaction that results in the release of SRP (Stroud and Walter, 1999; Keenan et al., 2001). Ribosomal components are critical for GTP cycling during SRP:SR interactions. First, GTP binding by SRP54 is required for interaction with SR and release of the nascent peptide at the ER membrane; a ribosomal constituent acts as a guanine-nucleotide exchange factor for the SRP54 GTPase by increasing the affinity of SRP54 for GTP (Bacher et al., 1996). A 21 kDa r-protein acts as a GTPase activating protein for the SR β subunit GTPase at the translocon, stimulating GTP hydrolysis and subsequently reducing the affinity of SR for both GTP and GDP (Bacher et al., 1999; Fulga et al., 2001). Binding of the ribosome to the translocon at the ER membrane displaces the SR from the r-protein (Fulga et al., 2001).

1.3.2.4. R- protein extraribosomal functions

In addition to their ribosomal roles, many r-proteins have extraribosomal functions, usually identified in the course of researching processes other than translation. A decade ago, Wool (1996 and references therein) reviewed more than 30 cases where r-proteins had shown involvement in processes such as replication and DNA repair (prokaryotic S1, L14, S9; eukaryotic S3, P0), transcription and RNA processing (prokaryotic S10, S12; eukaryotic S20), autogenous translational control (prokaryotic L1, L4, L10, S4, S7, S8; eukaryotic L4, L7, L30, S14), development (eukaryotic S2, S4, S6, S15a, S18, S19), and various eukaryotic cellular processes (cancer - L5; mitosis – S8; differentiation – L32; iron-distribution – P2). Numerous human r-protein genes map to candidate loci for congenital diseases, however, it is unlikely that all of these disorders, such as retinal degeneration, result solely from decreased ribosomal function

(Uechi et al., 2001). Diamond-Blackfan anemia, a human red blood cell (erythrocyte) disorder, is one such disease, and can result from a mutation in *S19* (Miyake et al., 2005). Suppression of *S19* expression results in a decrease in hematopoietic (erythrocyte progenitor) cell proliferation and differentiation of erythroid cells (Miyake et al., 2005). Mutation in *Drosophila S3* results in the *Minute* phenotype, a homozygous lethal mutation with numerous developmental abnormalities associated with heterozygotes (Andersson et al., 1994). In Arabidopsis, r-protein S18 appears to play a role in developmental regulation; a mutation in the copy of *S18* at the *PFL* locus causes pointed first leaves and growth retardation (Van Lijsebettens et al., 1994). Another Arabidopsis r-protein, L24, plays a role in gynoecium development, with mutants showing a stunted ovary. Studies of similar mutants suggest that L24 may play a role in reinitiation of translation of auxin response factor transcripts (Nishimura et al., 2004). Although it does not yet have a documented extraribosomal function, Arabidopsis *L7A* is transcribed in cell culture, but the protein is not assembled into ribosomes, suggesting it may have a role outside of translation (Chang et al., 2005).

Other r-proteins have been identified during screening for genes involved in specific cellular or stress processes. As a result, eukaryotic r-proteins L19 (Sonnemann et al., 1991), S4, S8, S14, S15a, L13a, and L22 (Shen et al., 2005) have all been identified as Ca^{2+}-dependent calmodulin binding proteins, indicating possible roles for these proteins in calmodulin-mediated signal transduction. Studies of eukaryotic cell growth, transformation (conversion to uncontrolled division), and death (apoptosis) have identified numerous r-proteins, including S3a (Naora and Naora, 1999) and L23 (Jin et al., 2004), to be involved in distinct aspects of cell cycle control. In Arabidopsis, S27a has been implicated in the rapid degradation of mRNA following DNA damage from chemical (methyl methane sulfate) or UV treatment. While S27a appears dispensable under normal growing conditions, mutants show hypersensitivity, tumor growth, and increased mRNA stability in response to these genotoxic stresses (Revenkova et al., 1999). R-protein genes in *E. coli* can be downregulated (*L2*, *L3*, *L4*, *S3*, *S5*, *S16*), or upregulated (*L9*, *L20*) following cadmium exposure and although the overall response involves translational repression, r-proteins are expressed differentially in order to achieve this result (Wang and Crowley, 2005).

Perhaps the best documented r-protein with extraribosomal functions is RACK1 (receptor of activated C-kinase), a well-characterized scaffold protein, and recently confirmed r-protein (Sengupta et al., 2004), that interacts with multiple kinases and membrane receptors as it coordinates signal transduction pathways (reviewed in Nilsson et al., 2004). Previously reported as a ribosome-associated protein (Link et al., 1999; Shor et al., 2003), cryo-EM of the yeast ribosome has established RACK1 as a 40S subunit protein, localized to the head of the 40S subunit where it interacts with the 18S 3' major domain near the mRNA exit site (near the platform on the solvent-exposed back of the 40S subunit; Sengupta et al., 2004). The presence and subunit association of RACK1 has also been confirmed in plant ribosomes (Chang et al., 2005). RACK1 interacts with numerous soluble signaling molecules (e.g. protein kinase C, Src kinase) as well as various membrane receptors (e.g. integrin receptor), possibly functioning to recruit signaling molecules to ribosomes and ribosomes to membranes for site-specific translation (Nilsson et al., 2004). RACK1 can also recruit specific mRNAs for translation, including mRNA encoding r-protein L25 in fission yeast (Shor et al., 2003).

At least three other r-proteins have documented extraribosomal roles in translation and transcription regulation. L30 (L30e), an archaebacterial and eukaryotic component of two intersubunit bridges (Halic et al., 2005), is able to regulate both pre-mRNA splicing in the nucleus and translation in the cytoplasm through partially induced-fit, negative autoregulatory binding interactions with its own transcript (Mao et al., 1999). Interestingly, L30 also has a role in the translational recoding of the UGA stop codon to selenocysteine in eukaryotic selenoprotein mRNAs, a process that involves the binding of the selenocysteine insertion sequence (SECIS) found in the 3' untranslated region (3' UTR) of selenoprotein transcripts (Chavatte et al., 2005). In its recognition of the SECIS (by adopting a kink-turn-like motif induced by Mg^{2+}), L30 demonstrates a remarkable RNA binding capability that allows it to play a role in diverse regulatory processes (Chavatte et al., 2005). In human cells, L13a has been identified as a translational regulator of ceruloplasmin (Cp), a protein upregulated by interferon-gamma (IFNγ) and expressed during host defence immune responses (Mazumder et al., 2003). Over an 8-24 h treatment of cell culture with IFNγ, L13a is phosphorylated and released from ribosomes. In its phosphorylated form, L13a inhibits

the translation of Cp mRNA by binding to the 3' UTR IFNγ-activated inhibitor of translation (GAIT) element of the Cp transcript (Mazumder et al., 2003).

In *E. coli*, a group of r-proteins (S4, L3, L4, and L13) are able to prevent termination of rRNA transcription (Torres et al., 2001). In particular, S4 is a general transcription antitermination factor, interacting with RNA polymerase directly and preventing transcription termination at Rho-dependent termination sites (Torres et al., 2001). Interestingly, like L30, S4 recognizes a unique RNA structure (pseudoknot) and autoregulates expression of its own operon; the antitermination factor activity of *E. coli* r-proteins in upregulating rRNA expression may act in concert with repression of r-protein operons in order to ensure stoichiometric amounts of ribosomal components (Torres et al., 2001).

1.3.2.4.1. R- proteins expand the functional reach of the ribosome

With r-proteins such as eukaryotic RACK1 and L13a and prokaryotic S4, it is clear that the ribosome, in addition to its core peptidyltransferase and decoding activities, is much like RNA polymerase, a complex of multiple functional components serving as a regulator of gene expression. The involvement of r-proteins in extraribosomal activities supports the theory that proteins were recruited to the ribosome from other functions in the cell over evolutionary time, thus adding to a rRNA core (Wool, 1996). Like transcription factors, associating with an RNA Pol II initiation complex to regulate specific gene expression, r-proteins have expanded the regulatory abilities of the ribosome. R-proteins can have positive regulatory effects, recruiting specific transcripts for translation (RACK1), or negative effects, inhibiting translation by leaving the ribosome (L13a, L30), effect the modification of nascent peptides through the translational machinery (e.g. L30 and selenoproteins), or by recruiting regulatory molecules (e.g. kinase recruitment by RACK1). Change in individual r-protein gene expression in response to stressors (e.g. metal stress, immune response) also indicates that r-proteins truly make the ribosome a multifunctional complex as a repository of different regulatory molecules.

1.4. Ribosomal protein genes

Given the structural and functional diversity of r-proteins, it is perhaps not surprising that there is also a great deal of diversity when it comes to the genes that

encode them. While r-protein genes in *E. coli* are arranged in 20 operons, with approximately half of the genes mapping to a single locus (Mager, 1988), most eukaryotic r-protein genes are dispersed throughout the genome (Mager, 1988; Barakat et al., 2001; Uechi et al., 2001) and are generally present in multiple copies (e.g. Wool et al., 1995; Barakat et al., 2001). Although eukaryotic r-protein genes are usually found in multi-gene families, the organization of those families differs widely between organisms. Plants, such as rice, maize, and Arabidopsis, have lower copy number families than organisms such as the rat (reviewed in Wu et al., 1995), but more plant r-protein gene family members tend to be functional. For example, of 59 characterized rat r-protein genes, the average copy number is 12 (Wool et al., 1995), but despite the large r-protein gene families in mammals, usually only one gene is functional and the rest are probable pseudogenes (Wool et al., 1995). For example, there are over 2,000 processed r-protein pseudogenes in the human genome (Harrison et al., 2002; Zhang et al., 2002). In contrast, plants tend to have multiple expressed r-protein genes from each family (Wu et al., 1995; Barakat et al., 2001). Arabidopsis r-protein genes are present in multigene families of 2 to 7 members with an average copy number of 3, and most members of each family are expressed (Barakat et al., 2001). Expression of multiple r-protein genes from a single family may be necessary to accommodate high translational needs in growing plant tissue; thus, r-protein gene copies under developmental regulation may be required in addition to those gene copies that are constitutively expressed (Van Lijsebettens et al., 1994). For example, Arabidopsis r-protein gene *L16* is present as two copies in the genome, with one isoform expressed in proliferating tissues while the other is expressed in more specific tissues (Williams and Sussex, 1995). Expression of multiple gene family members may also be indicative of multiple functions for r-proteins from any given gene family, with some members fulfilling ribosomal functions and others having extraribosomal roles. Alternatively, expression of multiple genes in an r-protein gene family may be indicative of ribosome heterogeneity (e.g. Chang et al., 2005; Giavalisco et al., 2005).

1.5. Ribosomal protein gene expression and regulation

The coordinated expression of ~50-80 r-proteins to produce equimolar amounts for each ribosome presents a great challenge to every cell. With the exception of the

acidic stalk proteins, r-proteins are generally found as a single copy per ribosome (Hardy, 1975; Subramanian, 1975; Tal et al., 1990); this has been confirmed in recent years by crystallographic studies of individual subunits (e.g. Wimberly et al., 2000; Nissen et al., 2000). Despite the relatively high overall conservation of amino acid sequence and structure between r-proteins, the primary mechanisms of r-protein gene regulation differ between species (reviewed in Mager, 1988). While bacterial r-protein genes are arranged in operons, the dispersal of eukaryotic r-protein genes across genomes and the presence of multiple expressed genes for each r-protein complicate coordinated regulation substantially. Not only are eukaryotic r-protein genes widely scattered between chromosomes, the probable recruitment of r-proteins to the ribosome from other functions over evolutionary time (e.g. Wool, 1996) also suggests that a diverse array of regulatory mechanisms for each r-protein gene may be expected, especially for non-universally conserved r-proteins (i.e. archaea/eukarya-specific, eukarya-specific).

1.5.1. Prokaryotic r-protein genes: operons and feedback inhibition

The control of r-protein gene expression in prokaryotes, primarily studied in *E. coli*, is elegant and streamlined, allowing rapid changes in the expression of ribosomal components in response to stimuli such as nutrient availability. Under stable growth conditions, coordinated synthesis of ribosomal components ensures ribosome biogenesis proportional to the growth rate, without significant accumulation of unincorporated ribosomal constituents (reviewed in Lindahl and Zengel, 1986; Nomura, 1999). During a nutritional upshift from poor to enriched growth medium or, conversely, under stringent conditions of nutritional limitation, synthesis of r-proteins (and rRNA) is rapidly up- or down-regulated, respectively (Lindahl and Zengel, 1986; Nomura, 1999).

The arrangement of prokaryotic r-protein genes in operons is a key part of the regulatory strategy for coordinate synthesis, allowing groups of r-protein genes to be regulated simultaneously. Among these operons, the *S10*, *spc*, and *alpha* gene clusters are the largest and most conserved. The *S10, spc, alpha* operons encode approximately half of the r-protein genes found in *E. coli* (Lindahl and Zengel, 1986; Mager, 1988; Coenye and Vandamme, 2005). The remaining r-protein genes are also found predominantly in operons, but in smaller clusters of one to four genes (Lindahl and

Zengel, 1986). The regulation of r-protein synthesis in *E. coli* primarily takes the form of assembly-mediated feedback control, where one of the r-proteins encoded by each operon serves as a negative regulator of that operon via transcriptional or, more commonly, translational inhibition ('autogenous regulation'; reviewed in Lindahl and Zengel, 1986; Nomura, 1999). In most cases, an r-protein binds to the polycistronic mRNA encoded by its operon, preventing translation of the r-proteins encoded in the message (e.g. Yates et al., 1980). Autogenous regulators (that are also known rRNA binders) have been identified for most of the r-protein-encoding operons in *E. coli*. For example, L1, S4, and S8 are translational repressors of the *L11*, *alpha*, and *spc* operons, respectively (Yates et al., 1980), S7 is a translational repressor of the *str* operon (Nomura et al., 1980), and L20 is an autogenous regulator of the *IF3* operon (Raibaud et al., 2003).

Initial reports of structural similarities between regulatory r-protein binding sites on rRNA and mRNA suggested a model for feedback regulation where the rRNA and mRNA binding sites compete for occupation by the r-protein (Nomura et al., 1980). This model has been supported by recent structural studies of autogenous regulator r-proteins in complex with cognate mRNA and rRNA binding sites. Crystal structures of S8 (Merianos et al., 2004) and L1 (Nevskaya et al., 2005) show that the interactions between these autogenous regulators and their 16S or 23S rRNA binding sites are highly similar to their interactions at their binding sites on mRNA due to similarities in RNA structures. NMR of L20 on its rRNA and mRNA binding sites also confirmed structural similarity between binding sites (Raibaud et al., 2003). For efficient autogenous regulation, r-proteins must have greater affinities for their rRNA binding sites than for their mRNA (Lindahl and Zengel, 1986); in agreement, L1 forms a more extensive network of contacts with its more complex rRNA binding site, than with its mRNA (Nevskaya et al., 2005). The higher affinity of S8 for rRNA than its mRNA also corresponds to a greater buried RNA surface area in the S8:rRNA interaction, where S8 binds to a 16S rRNA three-helix junction (Merianos et al., 2004).

Translational repression is not the only form of r-protein gene regulation in *E. coli*. In at least one case, the r-protein acting as a negative regulator causes termination of transcription of the operon (attenuation; e.g. Zengel et al., 1995; Stelzl et al., 2003).

L4 regulates both translation and transcription of the *S10* operon. It is able to bind and stimulate transcription termination at a site in the *S10* leader and is also able to bind and repress translation of the *S10* polycistronic mRNA (reviewed in Lindahl and Zengel, 1986; Stelzl et al., 2003). Binding assays between L4 and its *S10* leader mRNA and 23S rRNA sites confirm that these sites compete for L4 binding, although binding affinity between L4 and its 23S binding site is still 4-fold greater than between L4 and *S10* leader mRNA (Stelzl et al., 2003). In addition to autogenous regulation, control of r-protein gene expression involves other, less well understood, mechanisms at the level of transcription initiation and transcript stability (Lindahl and Zengel, 1986). Given the roles of *E. coli* r-proteins S4, L3, L4, and L13 in preventing the termination of rRNA transcription (Torres et al., 2001; see above, section 1.3.2.4, r-protein extraribosomal functions), it is clear that coordinated synthesis of ribosomal components in *E. coli* involves more than just autogenous regulation of r-protein operons.

While *E. coli* has long served as a model prokaryote, and the operon arrangement of r-protein genes is conserved, the extent to which the regulatory mechanisms controlling r-protein gene expression are conserved remains an open question. The *S10-spc-alpha* gene clusters are generally conserved among prokaryotes, although a recent survey of bacterial and archaebacterial genomes (99 bacterial species, 12 archaebacteria) found that gene order and content in the clusters has been subject to rearrangements, losses, and additions between species (Coenye and Vandamme, 2005). In addition to the similar r-protein gene arrangements between prokaryotic species, some experimental data confirms similar regulatory mechanisms between specific species. For example, L1 proteins from diverse bacterial and archaebacterial species have been shown to bind both heterologous rRNA and mRNA target sites, although not with the same affinity (Köhrer et al., 1998). A similar mechanism of *S10* operon regulation between diverse prokaryotes has been suggested by the ability of heterologous L4 from both gram-negative and gram-positive bacteria to repress *S10* expression in *E. coli* (Zengel et al., 1995). Another eubacterium, *Vibrio cholerae*, also shows L4-mediated autogenous control of the *S10* operon and S8-mediated regulation of the *spc* operon as in *E. coli* (Allen et al., 2004). *S10* leaders from other bacterial species have also demonstrated competence to respond to regulation by *E. coli* L4 (Allen et al., 1999). The *S10* operon

from the eubacterium *Pseudomonas aeruginosa*, however, is not L4-responsive, showing at least some diversity in r-protein gene regulation among prokaryotes (Allen et al., 1999).

1.5.2. Yeast r-protein genes: transcriptional regulation

Like prokaryotes, the unicellular fungus *Saccharomyces cerevisiae* (Baker's yeast) has an efficient system for coordinating r-protein gene expression in response to cellular needs. The r-protein genes of yeast account for 60% of the 30 most transcribed genes of the yeast transcriptome, with between 60 (RPS18B) and 207 (RPL10) mRNAs per cell, depending on the gene (Velculescu et al., 1997), and they constitute almost half of the yeast genes expressed at a rate of over 30 mRNAs per hour (Lee et al., 2004). With ~2000 ribosomes synthesized per minute in an actively growing yeast cell, r-protein gene expression involves a major portion of the transcription, post-transcriptional processing (~90% of all mRNA splicing), translation, and transport events in the cell (Warner, 1999). Even with such high levels of r-protein gene expression during growth, the yeast cell is an efficient regulator of ribosome biogenesis, with amounts of r-proteins and rRNAs generally not exceeding cellular requirements (Warner et al., 1985; Warner, 1989). Transcription of r-protein genes is upregulated in response to a nutritional upshift or active growth phase, and downregulated in response to heat stress, starvation (amino acid/nitrogen deprivation), or entry into stationary phase (reviewed in Warner, 1989; Li et al., 1999; Warner, 1999). Transcriptional activation or repression of r-protein genes can respond to changes in environmental conditions in 15 to 30 min or less (Li et al., 1999; Warner, 1999). Any defect in the secretory pathway (ER-Golgi-membrane transport) also results in a reduction of r-protein gene transcription (Mizuta and Warner, 1994). In addition to direct transcriptional control of r-protein gene expression, yeast r-protein mRNAs have short half-lives (typical $t_{1/2}$ estimated at 5-7 min; Li et al., 1999), and excess r-proteins are quickly degraded in response to overproduction (Warner et al., 1985), both ensuring the rapid turnover required for equimolar r-protein production in response to cellular requirements.

Unlike the case in prokaryotes (but as in other eukaryotes), the 137 genes that encode the 78 different yeast r-proteins (79 r-proteins including RACK1, Link et al., 1999) are dispersed throughout the genome (Planta and Mager, 1998). Interestingly, 59

of these r-proteins are encoded by functional duplicate genes, and over 100 of the 137 genes have introns (Planta and Mager, 1998). R-protein genes thus account for ~one third of the intron-containing genes of the entire *S. cerevisiae* genome, and tend to have the largest introns found in yeast genes, suggesting a role for introns in the regulation of r-protein gene expression (Spingola et al., 1999). While yeast r-proteins are often encoded by two functional genes, duplicate genes are not transcribed at the same level (Warner et al., 1985; Jiménez et al., 2002). There is evidence to suggest that duplicate r-protein genes may each be transcribed at half the rate of a single copy gene (Warner et al., 1985), but degradation of excess mRNAs and r-proteins may also be responsible for maintaining equimolarity of r-proteins produced by multi-copy genes (Warner, 1989). While there are rare cases of autogenous regulation of yeast r-protein genes (L2, Presutti et al., 1995; L30, Mao et al., 1999, Vilardell et al., 2000) via splicing inhibition, mRNA destabilization, and translation, coordinated yeast r-protein gene expression is primarily regulated at the level of transcription.

Although they are dispersed throughout the genome, yeast r-protein genes share a characteristic set of motifs in their upstream regulatory regions that facilitate coordinated transcription. Early studies identified duplicated upstream activation sequences (UAS), called RPG and HOMOL1 boxes, in the 5' regulatory regions of r-protein genes that were required for transcription; a T-rich region downstream of the UAS was found to enhance transcription (Rotenberg and Woolford, 1986; Woudt et al., 1986; Schwindinger and Warner, 1987; reviewed in Mager, 1988). A protein initially identified as "TUF" was found to bind to the UAS motifs (Vignais et al., 1987; reviewed in Warner, 1989), and was later identified as Rap1 (repressor-activator protein 1), a general yeast DNA binding factor (Warner, 1989).

The general architecture of yeast r-protein gene UAS/promoter regions consists of two Rap1 binding sites followed by a T-rich region; a few r-protein genes have an Abf1 (autonomously replicating sequence binding factor) binding site instead of two Rap1 sites (reviewed in Planta et al., 1995). Rap1 is an essential, multipurpose DNA binding protein that plays a role in both repression (e.g. silencing at mating loci, telomeres), and activation (e.g. glycolytic pathway genes, r-protein genes) of transcription. Both Rap1 and Abf1 displace nucleosomes at their binding sites, and

appear to play a role in chromatin reorganization, allowing promoter access to other regulatory factors (e.g. Sir proteins at silencing loci, Gcr1 at glycolytic pathway loci; reviewed in Planta et al., 1995). The spacing of duplicate Rap1 sites in r-protein gene promoters, found 200-500 bp upstream of the translation start codon, is very specific, with the motifs being separated by only 15-26 bp (Lascaris et al., 1999). Chromatin immunoprecipitation (ChIP; IP following crosslinking of proteins to DNA and chromatin shearing), using actively growing yeast, identified 122 r-protein gene promoters enriched in Rap1, including genes with non-canonical Rap1 binding sites (Lieb et al., 2001). Probing of yeast genome microarrays with the Rap1-enriched fragments identified the Rap1-bound genes as the most actively transcribed in the genome (Lieb et al., 2001). Abf1 also confers high-level transcription in concert with the T-rich region; Abf1 reorganizes the chromatin structure up- and downstream of its binding site, and the T-rich sequence maintains a nucleosome-free region downstream of Abf1 (Lascaris et al., 2000).

While Rap1 and Abf1 are DNA binding proteins that bind during activation or repression, recent research has sought to identify other factors that bind to yeast r-protein promoters to confer specificity of expression, such as in response to changes in growth or environmental conditions. Histone acetylation is associated with activation of transcription, and Esa1, a histone acetylase, has been shown to be recruited to r-protein gene promoters in a Rap1- and Abf1-dependent manner, resulting in acetylation of histones associated with the r-protein genes (Reid et al., 2000). Following amino acid starvation or a heat shock, Esa1 occupancy on r-protein gene promoters is depleted, histone acetylation is decreased, and r-protein gene transcription is repressed; Rap1 occupancy, however, remains constant, suggesting that recruitment of Esa1 is associated with gene activation (Reid et al., 2000). This suggestion is supported by the fact that Esa1 is not recruited to Rap1-bound telomeric regions (Reid et al., 2000), and the independent observation that following a heat shock, r-protein genes (which are nucleosome-depeleted when active) are repressed and nucleosome occupancy increases, but Rap1 remains bound to r-protein gene promoters (Lee et al., 2004).

The TOR (target of rapamycin) pathway in yeast, mediated by the serine/threonine kinases TOR1 and TOR2, is a signal transduction pathway that

regulates biosynthesis by the cell, including activation of r-protein gene expression, in response to nutrient abundance (Warner, 1999; Powers and Walter, 1999). The drug rapamycin inhibits the TOR pathway, resulting in transcriptional repression of r-protein genes even during a nutritional upshift, indicating the importance of the TOR pathway in transducing nutrient signals to r-protein genes (Powers and Walter, 1999). TOR signaling is required to maintain Esa1 occupation at r-protein genes, and r-protein gene repression as a result of rapamycin treatment or nutrient depletion involves release of Esa1 and occupation of r-protein genes by the Rpd3-Sin3 histone deacetylase complex (Rohde and Cardenas, 2003; Figure 1.4).

A transcription factor with a forkhead DNA binding domain, Fhl1 (forkhead-like), found specifically at r-protein gene promoters (Lee et al., 2002), has recently been implicated in TOR-mediated (and TOR-independent) transcriptional regulation of r-protein genes. The association of Fhl1 and its cofactors (see below; Figure 1.4) with r-protein gene promoters is enhanced by interaction with Rap1 (Wade et al., 2004) and does not appear to involve direct DNA binding (Rudra et al., 2005). Under normal growth conditions, in cells not treated with rapamycin, Fhl1 is associated with r-protein genes and binds a coactivator, Ifh1 (interacting with forkhead 1), activating r-protein gene transcription (Martin et al., 2004; Schawalder et al., 2004; Wade et al., 2004; Rudra et al., 2005). During repression of r-protein gene transcription as a result of stress (Wade et al., 2004), rapamycin treatment (Martin et al., 2004; Schawalder et al., 2004; Wade et al., 2004; Rudra et al., 2005), or nutrient deprivation (Martin et al., 2004; Schawalder et al., 2004), Ifh1 (but not Fhl1) dissociates from r-protein gene promoters. While Fhl1 remains bound to r-protein gene promoters during rapamycin treatment or nutrient depletion, it binds a corepressor, Crf1 (corepressor with Fhl1); both Ifh1 and Crf1 require direct interaction with the same domain of Fhl1 in order to associate with r-protein genes (Martin et al., 2004). Crf1, unlike the nuclear-localized Ifh1, is retained in the cytoplasm by a TOR-mediated protein kinase A (PKA) interaction; upon repression of TOR by rapamycin or nutrient deprivation, Crf1 is phosphorylated and accumulates in the nucleus where it can compete with Ifh1 to bind Fhl1 and repress r-protein genes (Martin et al., 2004). Despite this elegant system, the low-level of r-protein gene transcription that can occur without Fhl1 and Ifh1 can still be repressed

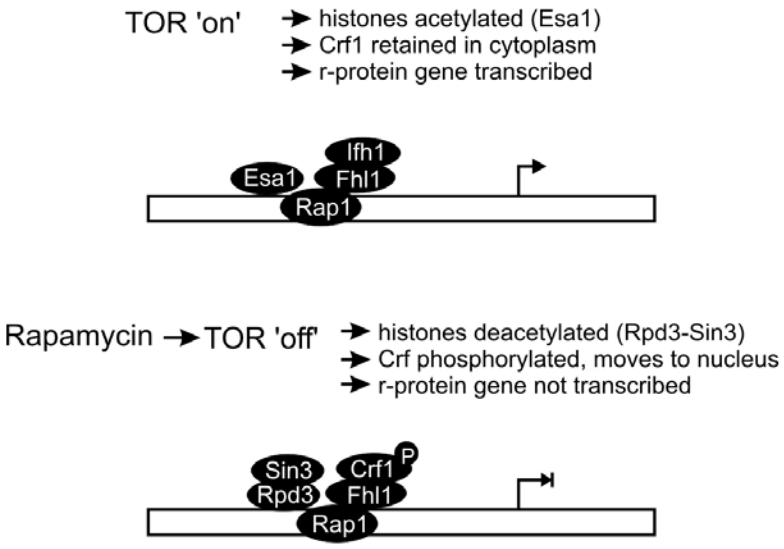


Figure 1.4. Schematic of TOR-mediated factors activating or repressing yeast r-protein gene transcription. Rap1 displaces nucleosomes, enabling association of histone acetylases (Esa1) or deacetylases (Rpd3-Sin3 complex), and transcription factors (Fhl1). Under activation conditions, Fhl1 coactivator Ifh1 binds, and corepressor Crf1 is maintained in the cytoplasm in dephosphorylated form. Under deactivation conditions, Crf is phosphorylated and moves into the nucleus, displacing Ifh1 and resulting in transcriptional repression of the r-protein gene. See text and Martin et al. (2004) for details.

by rapamycin, indicating that even more regulatory factors play a role in this pathway (Rudra et al., 2005).

1.5.3. Animal r-protein genes: 5' TOP mRNAs and translational regulation

The separation of transcriptional and translational machinery by the nucleus in all eukaryotes complicates the regulation of r-protein gene expression; this is further complicated in multicellular organisms by the challenges of responding to tissue-specific, developmental, and environmental signals generated through multiple, complex transduction pathways. Despite the detailed knowledge of r-protein gene regulation in *E. coli* and yeast, control of coordinated r-protein expression in animals and plants remains a fundamentally unsolved problem. In animals, the best-studied systems are primarily mammalian (i.e. rat, mouse, human), although some research has been carried out using the amphibian *Xenopus laevis*. Regulation of r-protein gene expression in these systems has been observed at the transcriptional, post-transcriptional, and translational levels, with translational regulation playing a primary role (Mager, 1988; Larson et al., 1991).

As in other organisms, r-protein gene expression in animals is closely linked to growth and development, with high levels of expression in mitotically active, growing and/or developing tissues and low levels of expression in mature or quiescent tissues. During embryogenesis in Xenopus, r-protein mRNA levels increase rapidly between gastrulation and neurulation and are then stored as mRNPs (mRNA-ribonucleoprotein particles) in the cytoplasm; several hours following mRNA synthesis, r-protein mRNAs are actively translated in the embryo (Pierandrei-Amaldi et al., 1982; reviewed in Amaldi et al., 1989). Active r-protein gene expression (transcription, translation) has also been documented in mouse fibroblast culture following a growth stimulus (Geyer et al., 1982), mouse spermatogenic tissues (testes, spermatocytes, spermatids; Kleene et al., 2003), and during rat liver development and regeneration following injury (Aloni et al., 1992; Chung et al., 2005). In contrast, during differentiation from rapidly-dividing myoblasts to non-dividing muscle fibers in mouse cell culture, r-protein gene expression is reduced 2-fold as a result of decreased transcription and translation (Agrawal and Bowman, 1987).

Animal r-protein genes are primarily regulated at the translational level, with their mRNAs alternating between repression by sequestration in mRNPs (also called the

subpolysomal fraction), and active expression in translating polysomes (reviewed in Meyuhas, 2000). As mentioned above, there is a delay of several hours between r-protein gene transcription and translation during embryogenesis in Xenopus (Pierandrei-Amaldi et al., 1982; Amaldi et al., 1989; Amaldi et al., 1995). Corresponding to this delay, r-protein mRNAs in the early stages of embryo development are present only in the subpolysomal fraction; later embryonic stages contain a majority of r-protein mRNAs associated with polysomes (Pierandrei-Amaldi et al., 1982). This change in distribution of r-protein mRNAs from mRNPs to polysomes is also seen during a nutritional upshift following serum starvation of Xenopus kidney cells (Loreni and Amaldi, 1992). The same pattern of translational regulation has been observed in mammals, with ~70-85% of r-protein mRNAs associated with polysomes in active cells but only ~20-50% of r-protein mRNAs in the polysomal fraction of resting or mature cells (Geyer et al., 1982; Aloni et al., 1992). Unlike the short half-life of yeast r-protein mRNAs, animal r-protein mRNAs have a much slower turnover rate, with half-lives of over 8 h (Geyer et al., 1982). r-protein mRNAs are poorly translated in the absence of rRNA synthesis or in the presence of a full complement of ribosomes (Pierandrei-Amaldi et al., 1982; Amaldi et al., 1989; Amaldi et al., 1995) and excess r-proteins are degraded quite rapidly ($t_{1/2}$ ~1 h in mutant embryos lacking rRNA synthesis; Pierandrei-Amaldi et al., 1985).

A key feature of virtually all vertebrate, and at least some invertebrate, r-protein mRNAs for translational regulation is the 5' terminal oligopyrimidine tract (5' TOP) located in the 5' UTRs (Meyuhas, 2000). Transcripts start with the 5' TOP motif: a C residue at the cap site followed by a stretch of pyrimidines (e.g. 8-12 pyrimidines in Xenopus, Amaldi et al., 1995; 4-14 in mammals, Meyuhas, 2000) which is then followed by GC-rich sequence (Amaldi et al., 1995; Meyuhas, 2000). The 5' UTRs of animal r-protein mRNAs tend to be short, 35-50 nt in Xenopus (Amaldi et al., 1995), ~20-50 nt in mammals (Meyuhas and Klein, 1990; Meyuhas, 2000). Likewise, the 3' UTRs of vertebrate r-protein mRNAs are also very short, for example, 40-60 nt in Xenopus (Amaldi et al, 1995) and an average of 35 nt (range of 5-250 nt) in human (20 times shorter than in most other human mRNAs; Ledda et al., 2005). Under conditions of serum starvation in human cell culture, the stringency of translational repression of

r-protein mRNAs is dependent on 3' UTR length in addition to the TOP 5' UTR (Ledda et al., 2005). Interestingly, 3' UTR length does not appear to affect r-protein mRNA stability, and transcripts are stable for over 10 h in dividing cells (Ledda et al., 2005).

Even though the importance of the 5' TOP motif in translation has been established (Meyuhas, 2000), it is not yet fully understood how translational control of r-protein mRNAs is achieved or what specific role the 5' TOP plays. Changes in translational efficiency of vertebrate r-protein mRNAs as measured by association with polysomes has shown a direct correlation with changes in eukaryotic initiation factor eIF-4E mRNA abundance (Aloni et al., 1992) and in phosphorylation of eIF-4E and r-protein S6 (reviewed in Stewart and Thomas, 1994; Amaldi et al., 1995). As in yeast, the mammalian target of rapamycin (mTOR, the mammalian homologue of the yeast TOR proteins) pathway has been implicated in animal r-protein gene expression, but via translational control rather than transcription. Recently, rapamycin was shown to repress translation of r-protein mRNAs *in vivo* during starvation recovery in rat liver (Reiter et al., 2004), possibly by blocking the phosphorylation of both S6 and the primary S6 kinase ($p70^{s6k}$) and interfering with 5' TOP mRNA mobilization into polysomes (Stewart and Thomas, 1994). More recent work has identified a complex containing mTOR that is responsible for phosphorylating both $p70^{s6k}$ and eIF-4E binding protein (4E-BP1) and stimulating protein synthesis (reviewed in Martin and Hall, 2005). mTORC1 (mTOR complex 1; contains mTOR and mTOR-associated proteins mLST8 and raptor) activates $p70^{s6k}$ (and thus S6) and derepresses eIF-4E via phosphorylation of 4E-BP1. mTORC1 responds to amino acid and growth factor stimuli via two related signaling pathways involving numerous protein kinases (i.e. protein kinase B/PKB, phosphoinositide-dependent kinase 1/PDK1, phosphatidylinositol 3-kinase/PI3K) and a small GTPase (Rheb; reviewed in Martin and Hall, 2005).

Although TOR plays a key role in yeast and mammal r-protein gene regulation, the mechanisms of the pathway are very different. Phosphorylation (and dephosphorylation) of the yeast S6 homologue has no effect on cell growth, carbon source utilization, or heat shock sensitivity, suggesting that S6 phosphorylation has no role in yeast r-protein gene regulation (Johnson and Warner, 1987). The primary r-protein gene regulatory mechanisms, transcription versus translation, also differ

between yeast and animals, and a key question remains as to how S6 phosphorylation stimulates selective translation of r-protein mRNAs in animal systems. The search for factors that interact directly with r-protein 5' TOP mRNAs has not yielded many promising results thus far (Meyuhas, 2000). A recent study, however, has identified three factors (tumor suppressor protein p53, metal response element-binding transcription factor MTF-1 and phosphoprotein La autoantigen) associated with S25 mRNA in the nuclei of amino acid-starved hepatoma cells but not in nourished cells (Adilakshmi and Laine, 2002). An inhibition of a hnRNP A1 (heterogeneous nuclear RNP A1; acts as a mRNA shuttle between the nucleus and cytoplasm) association with *S25* mRNA by p53, La, and MTF-1, results in the sequestration of *S25* mRNA in the nuclei of starved cells (Adilakshmi and Laine, 2002).

Animal r-protein genes share a number of features in their regulatory regions, in addition to the common architecture of the 5' and 3' UTRs of their mRNAs. These conserved features were recently catalogued for mammals in a survey of the entire set of mouse and human r-protein genes (Perry, 2005). Contrary to the previous assumption that most animal r-protein genes do not contain TATA boxes (Mager, 1988), the Perry (2005) survey of mammalian r-protein genes determined that 60% of the genes have a canonical or non-canonical (A/T-rich) TATA box at the expected -25 position. The survey of r-protein gene transcription start sites, known to be located within polypyrimidine tracts in other vertebrates (Mager, 1988; Meyuhas and Klein, 1990; Amaldi et al., 1995; see above), yielded a consensus initiator sequence: $(Y)_2C^{+1}TY(T)_2(Y)_3$ (Perry, 2005). In general, r-protein gene upstream regulatory regions are short, showing sequence conservation between human and mouse only within ~200 bp of the transcription start site (Perry, 2005). Consensus motifs for the transcription factors YY1 (Yin Yang 1) and GABP (GA-binding protein) are each found in over 50% of mammalian r-protein gene regulatory regions and are common in the r-protein genes of other vertebrates. Like Rap1 and Abf1 in yeast, GABP (a heterodimeric transcription factor; Genuario and Perry, 1996) and YY1 (see review in Thomas and Seto, 1999) can act as transcriptional activators or repressors. GABP complexes bind to two sites near the transcription start site of the mouse *S16* gene, inhibiting its transcription, whereas GABP complexes bind and stimulate transcription of the *L30* gene (Genuario and Perry,

1996). GABP and YY1 gain additional specificity of function via interactions with other factors; YY1 interacts with numerous other proteins, including histone acetyltransferases and histone deacetylases (Thomas and Seto, 1999). GABP and YY1 binding motifs are generally absent from the promoters of other TOP-containing genes, indicating some *cis* element specificity for r-protein regulation (Perry, 2005). Such specificity could account for the lack of shared regulatory motifs between r-protein genes and other RNA Pol II-transcribed genes described in early studies (Zahradka et al., 1990).

Interestingly, almost half (47%) of mammalian r-protein genes have an intron in the 5' regulatory region, and the translation start codon is at the 3'-most end of exon 1 in another 29% of the genes (Perry, 2005). This feature has also been noted in Xenopus r-protein genes where the first intron is found immediately (or close to) following the ATG start codon, effectively separating the coding region from the upstream regulatory region (Amaldi et al., 1995). Indeed, a number of regulatory elements have been identified in the first introns of mammalian *L7* (Meyuhas and Klein, 1990), *L32* (Chung and Perry, 1989; Chung and Perry, 1993), and *S14* (Tasheva and Roufa, 1995) r-protein genes. In a unique regulatory mechanism, the first intron of human *S14* encodes two antisense RNA fragments that were found to interact specifically with *S14* and stimulate transcription *in vitro* and *in vivo* (Tasheva and Roufa, 1995). More conventional regulation is provided by the first introns of *L7* (Meyuhas and Klein, 1990) and *L32* (Chung and Perry, 1993), which harbor motifs for YY1 (also called δ; see above) at their 5' ends.

In addition to harboring *cis* regulatory elements, introns play a critical role in the post-transcriptional regulation of some animal r-protein genes, either by yielding alternative transcripts or by providing a target for mRNA degradation. Alternative splicing of mouse *S24* mRNA yields three different transcripts (a, b, and c) that demonstrate cell-specific abundance during adipocyte differentiation and tumorigenic transformation; for example, levels of S24(a) and (b) remain steady while S24(c) levels are down-regulated during adipocyte maturation (Xu et al., 1994). Alternative or inefficient splicing can also lead to a reduced abundance of r-protein gene transcripts. During development and regeneration in rat liver (Aloni et al., 1992) and embryogenesis in Xenopus (Amaldi et al., 1989), a decline in abundance of r-protein mRNAs was

determined to result from a post-transcriptional mechanism. The inefficient, incomplete, or lack of processing (in mRNAs transcribed from mutant genes lacking introns) of r-protein mRNAs has been identified as the cause of decreased r-protein transcripts for *L1* in Xenopus (Amaldi et al., 1989), *L32* in mouse (Chung and Perry, 1989), and at least four r-protein mRNAs in the nematode *Caenorhabditis elegans* (Mitrovich and Anderson, 2000). In *C. elegans*, alternative splicing yields transcripts for *L3*, *L7a*, *L10a*, and *L12* that contain premature stop codons, targeting the mRNAs for degradation by the 'mRNA surveillance' machinery (Mitrovich and Anderson, 2000). The introns involved in the alternative splicing of the *C. elegans* r-protein transcripts are highly conserved between nematode species. As with L2 (Presutti et al., 1995) and L30 (Mao et al., 1999, Vilardell et al., 2000) in yeast, some cases of post-transcriptional autoregulation have been identified in animal r-protein mRNAs. L1 in Xenopus (Amaldi et al., 1989) and L12 in *C. elegans* (Mitrovich and Anderson, 2000) can bind to their own transcripts, interfering with splicing and leading to a decrease in their own mRNA. Human S14 can bind its own mRNA and the antisense RNAs encoded in its first intron, leading to a decrease in *S14* transcripts (Tasheva and Roufa, 1995).

1.5.4. Plant r-protein genes: a new frontier

In contrast to the many relatively well-defined r-protein gene regulatory mechanisms in bacteria, yeast, and animals, very little is known about the regulation of expression of r-protein genes in plants. The r-protein genes of a few plants, primarily Arabidopsis, have been catalogued as a result of genome sequencing efforts (e.g. Barakat et al., 2001), and expression patterns of several individual r-protein genes have been determined, primarily at the transcriptional level (see below, section 1.5.4.1, transcriptional regulation). What is missing is an overall understanding of how the growth, developmental, and stress stimuli to which plant r-protein genes respond coordinates expression; indeed, the predominant mechanism of r-protein regulation in plants has yet to be identified. As seen in other organisms, the equimolar amounts of each r-protein in the ribosome are a result of the sum total of all regulatory mechanisms, i.e. transcriptional, post-transcriptional, translational, post-translational, on each r-protein gene. The unique feature of plant r-protein genes that complicates coordinate regulation is the presence of families of expressed genes for each r-protein (e.g. Wu et

al., 1995; Barakat et al., 2001). The following is a summary of current knowledge concerning the regulation of plant r-protein gene expression.

1.5.4.1. Transcriptional regulation

Past focus on plant r-protein gene expression has identified changes in transcript levels during development and/or following various stress or hormone treatments. R-protein genes in Arabidopsis (Van Lijsebettens et al., 1994; Williams and Sussex, 1995; Hulm et al., 2005; McIntosh and Bonham-Smith, 2005), *Brassica napus* (Bonham-Smith et al., 1992), potato (*Solanum tuberosum*; Taylor et al., 1992), tobacco (*Nicotiana tabacum*; Marty and Meyer, 1992; Gao et al., 1994; Dai et al., 1996; Popescu and Tumer, 2004), maize (*Zea mays*; Larkin et al., 1989; Lebrun and Freyssinet, 1991; Joanin et al., 1993; Chevalier et al., 1996; Dresselhaus et al., 1999; Williams et al., 2003), pea (*Pisum sativum*; Strafstrom and Sussex, 1992; Moran, 2000), petunia (*Petunia hybrida*; Lee et al., 1999), peach (*Prunus persica*; Giannino et al., 2000), and lupine (*Lupinus luteus*; Cherepneva et al., 2003) are expressed in almost all tissues that have been examined, with increased expression in meristems and other actively growing tissues. Transcripts for r-protein genes *S19* and *L7* in potato have been shown to increase in all plant tissues during induction of tuberization and then decrease as tuberization progresses, but with transcript levels lower in all tissue types compared to induced stolon tips, suggesting a developmental regulation of expression (Taylor et al., 1992).

Investigation of the developmental regulation of expression of Arabidopsis *L16* (*RPL11* according to the nomenclature of Barakat et al., 2001) also revealed specificity between members of this gene family (Williams and Sussex, 1995). While the upstream regulatory region of *RPL16B* drove reporter gene expression in vegetative as well as floral meristematic tissues, the *RPL16A* 5' regulatory region only conferred expression in root (lateral primordia, stele, behind the meristem), developing anthers (eventually restricted to the tapetum), and developing pollen (Williams and Sussex, 1995). In maize, at least two members of the *S21* r-protein gene family, A and B, are differentially regulated, such that *S21A* is highly expressed in eggs and zygotes, but shows only low expression in somatic tissues, while *S21B* is expressed in all tissues except eggs and zygotes (Dresselhaus et al., 1999). Different members of Arabidopsis r-protein gene families also show widely divergent expression patterns both in terms of transcript level

across tissue types, and in terms of timing of expression, during root development (Hughes and Friedman, 2005).

Mechanical wounding, auxin, and cytokinin treatments all result in an increase in transcript levels for many r-protein genes, with auxin in particular having a powerful effect on r-protein transcript levels (Gantt and Key, 1983, 1985; Gao et al., 1994; Van Lijsebettens et al., 1994; Dai et al., 1996; Cherepneva et al., 2003; Hulm et al., 2005; McIntosh and Bonham-Smith, 2005). For example, treatment of soybean seedlings with the synthetic auxin (2,4-dichlorophenoxy)acetic acid (2,4-D) resulted in an eight-fold increase in the level of r-protein mRNAs (Gantt and Key; 1983), and can also induce preferential transcription (Key et al., 1966) and processing (Melanson and Ingle, 1978) of rRNA. However, in contrast, treatment of pea seedling axillary buds with high concentrations of indole-3-acetic acid (IAA) repressed both bud growth and *L27* gene expression (Strafstrom and Sussex, 1992). The stress hormone abscisic acid (ABA) causes a decrease in r-protein transcript levels in species as distantly related as lupine (Cherepneva et al., 2003) and Arabidopsis (Hulm et al., 2005; McIntosh and Bonham-Smith, 2005).

Given the sessile nature of plants and their need to respond physiologically to environmental changes, it is perhaps not surprising that a number of stresses have also been found to affect r-protein gene transcript levels. For example, transcription of numerous r-protein genes has been shown to increase in maize in response to UV-B radiation (Casati and Walbot, 2003). Transcript levels for r-protein L13 from *B. napus* (Sáez-Vásquez et al., 2000) and S13, S6, and L37 from soybean (*Glycine max*; Kim et al., 2004) have all been shown to increase in response to cold treatment while transcript levels for r-protein S7 from rye (*Secale cereale*) decrease in response to cold stress (Berberich et al., 2000). R-protein L2 in soybean appears to be specifically regulated in response to biotic stress; *L2* transcript levels (from at least two genes) decrease during the hypersensitive response that results from a pathogen attack (Ludwig and Tenhaken, 2001). Treatment of soybean cell suspension cultures with abiotic pathogen-mimic elicitors (copper sulfate, glutathione, H_2O_2) also downregulates expression of the L2 genes (Ludwig and Tenhaken, 2001).

Transcription of multiple r-protein genes in response to growth, developmental, and stress stimuli indicates that plant r-protein genes are regulated, at least to some extent, at the transcriptional level. While general patterns of transcript abundance have been well-characterized for several plant cytoplasmic r-protein genes (see above), it has yet to be established how expression of the genes for the 75-92 cytoplasmic r-proteins in a plant, assumed to be a coordinated process (Browning, 1996), is achieved. The presence of multiple expressed genes encoding each r-protein complicates the coordination process, especially considering the evidence that gene family members often have divergent expression patterns (Williams and Sussex, 1995; Dresselhaus et al., 1999; Hughes and Friedman, 2005). This complexity is well-illustrated by L3 in tobacco, which is encoded by two genes: *RPL3A* and *RPL3B*. Knockouts of *RPL3A* result in an increase in *RPL3B* mRNA, yet RPL3B cannot fully compensate for the loss of *RPL3A* expression, and knockout plants show an altered phenotype with large mottled leaves showing an increase in leaf epidermal cell division and decrease in cell size (Popescu and Tumer, 2004). In contrast, repression of both *RPL3A* and *B* transcript levels decreases leaf epidermal cell division and increases cell size (Popescu and Tumer, 2004). Clearly the tobacco L3 isoforms display at least some functional divergence as well as complex coordinate regulation.

The few studies that have identified functional regulatory regions of plant r-protein genes have shown that, as in other organisms, the 5' regulatory regions tend to be short, conferring full transcriptional activity within a few hundred base pairs of the coding region (Dai et al., 1996; Moran, 2000). Putative regulatory motifs have been identified in some individual r-protein genes, such as the TATA boxes and a possible pollen-specific ("PB core") motif (*RPL16A* only) identified in the Arabidopsis *RPL16* genes (Williams and Sussex, 1995). A series of 5' deletions of the tobacco *L34* upstream flanking region identified a 50 bp region crucial for gene expression and a region further upstream important as a general enhancer (Dai et al., 1996). Interestingly, although auxin and wounding were shown to enhance *L34* transcription, no known auxin- (i.e. auxin-responsive elements/AuxREs; Guilfoyle et al., 1998) or wound-responsive *cis*-elements were found in the *L34* upstream regulatory region (Dai et al., 1996). A comparison of the upstream regulatory region from pea *L9* with >20 other plant r-protein

genes identified a putative regulatory motif, $^{5'}$TTAGGGTTTT$^{3'}$, in both forward and reverse orientations within 55 bp of the putative TATA box in most of the genes (Moran, 2000). The functional significance of the *L9* sequence, however, was not established by Moran; this motif (more frequently noted as $^{5'}$AAACCCTAA$^{3'}$) is known as the *telo* box, previously identified by other authors as present in numerous plant r-protein genes (Lenvik et al., 1994; Trémousaygue et al., 1999).

The *telo* box, *tef* box, and site II motif are the best characterized *cis* elements in plant r-protein genes, providing a link to r-protein gene expression in root meristems and other mitotically active tissues. The *telo* box is similar in sequence to the Arabidopsis telomere repeat sequence (Richards and Ausubel, 1988), and was subsequently identified as a "plant interstitial telomere motif," or "*telo* box", in the upstream regulatory regions of r-protein and other translation-related genes (Trémousaygue et al., 1999). The *telo* box motif (consensus $^{5'}$AAACCCTA$^{3'}$) has now been identified in 174 of 216 annotated Arabidopsis r-protein gene upstream regulatory regions (Trémousaygue et al., 2003), and drives gene expression, in concert with other *cis* elements, in root primordia (Manevski et al., 2000). The *telo* box alone does not confer specificity of expression, and must act with other elements (*tef* box; acidic r-protein *trap40* box) to drive expression in root meristems (Manevski et al., 2000).

The *tef* box, initially identified in the Arabidopsis elongation factor EF-1α A1 gene (Curie et al., 1991), has also been identified in plant r-protein and other genes, such as tobacco *thioredoxin h*, expressed in mitotically active, cycling cells (Regad et al., 1995; Manevski et al., 1999). The *tef* box motif (consensus $^{5'}$ARGGRYAnnnnnGTM$^{3'}$ where R = any purine, Y = any pyrimidine, and M = A or C) is usually associated with a *telo* box, but unlike the *telo* box, the *tef* box alone can confer specific expression, activating transcription in cells entering the cell cycle, undergoing the transition from quiescent to mitotically active states (Regad et al., 1995). Another *cis* element usually found in conjunction with the *telo* box is the *PCNA* (*PROLIFERATING CELL NUCLEAR ANTIGEN)* site II motif ($^{5'}$TGGGCC/T$^{3'}$), identified in 153 of the 174 *telo* box-containing Arabidopsis r-protein gene regulatory regions, almost exclusively present upstream of the *telo* box, which can be located up-, down- or both up- and downstream of the transcription start site (Trémousaygue et al., 2003). This same

arrangement of site II motifs upstream of *telo* boxes has also been found in 60 rice r-protein genes (Trémousaygue et al., 2003). Interestingly, the site II-*telo* box motif combination has also been identified in genes (r-protein, cell cycle) upregulated during axillary bud outgrowth following stem decapitation in Arabidopsis (Tatematsu et al., 2005), a developmental stage (bud outgrowth) that upregulates r-protein gene transcript levels in pea (Strafstrom and Sussex, 1992). The site II motif directs expression in actively dividing tissues, notably young leaves and root primordia, and the *telo* box enhances this expression (Trémousaygue et al., 2003).

Like the general transcription factor binding motifs found in r-protein genes of other organisms, the *telo* and *tef* boxes in plants are not r-protein gene specific. *telo* boxes are found in non-r-protein genes associated with mitotically active tissues such as *PCNA* (Manevski et al., 2000) and *tef* boxes regulate expression of a variety of genes associated with the cell cycle and redox regulation (Regad et al., 1995). Undoubtedly, specificity of expression for each r-protein gene is a matter of combinatorial control directed by multiple factors. The *tef* box itself has been found to nucleate formation of two different multimeric complexes, C1 and C2, *in vitro* using extracts from cycling Arabidopsis suspension culture cells (Manevski et al., 1999). C1 and C2 interact with different variations of the *tef* box consensus sequence, and while C1 can also interact with the acidic r-protein *trap40* motif, C2 does not interact with *trap40* and cannot form when the consensus *tef* motif ends in CC rather than AA (Manevski et al., 1999). Although C1 and C2 have some component factors in common (Manevski et al., 1999), clearly the multimeric nature of the complexes allows context-specific expression. In the case of *telo* box and site II-containing genes, site II is a target sequence for the transcription factor AtTCP20, a TCP (*TEOSINTE BRANCHED1*, *CYCLOIDEA*, *PCF*) domain protein (Trémousaygue et al., 2003). AtTCP20 was shown to interact directly with the *telo* box-binding (Trémousaygue et al., 1999) protein AtPurα in a yeast two-hybrid screen (Trémousaygue et al., 2003). The helix destabilizing ability of the human AtPurα homologue (Purα; Darbinian et al., 2001) suggests that, like yeast Rap1, Purα may serve as the basis for the recruitment of specific factors, including TCP20, to *telo* boxes and associated motifs such as the site II element (Trémousaygue et al., 2003; Tatematsu et al., 2005).

1.5.4.2. Post-transcriptional and translational regulation

In contrast to the r-protein genes of animals, there are little data available on the post-transcriptional or translational regulation of plant r-protein genes. The paucity of information does not necessarily reflect a lack of gene control at these levels, but probably reflects the lack of concerted research efforts on plant r-protein genes. While the transcript levels of plant r-protein genes have been noted in numerous cDNA library and microarray screening efforts, during the investigation of developmental and stress response profiles, few studies have focused on the r-protein genes themselves, and so many levels of regulation have been ignored.

One indication that post-transcriptional regulation may play a key role in plant r-protein gene control, as it does in other organisms, is the presence of r-protein gene introns in the otherwise intron-poor algal genomes. A survey of algal genomics noted that, of the ~464 genes of the *Guillardia theta* nucleomorph (red algal symbiont) genome, 17 contain introns, and 11 of those 17 intron-containing genes encode r-proteins (Grossman, 2005). As with other r-protein genes, the introns of the *G. theta* nucleomorph genes are located near the 5' ends of the primary transcripts (Grossman, 2005). Interestingly, transcripts from the two genes encoding peach S28 have also been found to contain introns between the 5' UTR and the ATG start codon, just as in animal r-protein transcripts (Giannino et al., 2000). At least two Arabidopsis r-protein gene familes, *S15a* and *L23a*, also contain introns upstream of their coding regions (K.B. McIntosh, J.L. Hulm, and P.C. Bonham-Smith, unpublished data; see Chapter 4, this volume). A full survey of plant r-protein gene structures, including exon/intron composition, has yet to be published. Given the propensity for r-protein genes to contain introns even in intron-poor genomes such as in yeast or algae, however, the presence of 5'-biased first introns is one of the first features that should be considered when identifying the probability of post-transcriptional regulation in plant r-protein genes. Peach *S28* mRNAs have been found in both spliced and probable unspliced forms, the former located in developing leaves and roots, and the latter found in fully expanded leaves, internodes and mature stems (Giannino et al., 2000). The presence of only mature *S28* mRNAs in mitotically active tissues indicates post-transcriptional

regulation of the *S28* gene, most likely due to accumulation of unspliced, inactive transcript (Giannino et al., 2000).

The storage of r-protein mRNAs in maize embryos for active translation upon germination (Beltrán-Peña et al., 1995) indicates that there is some translational regulation of r-protein gene expression in plants. Like the r-protein genes of animals, the 5' and 3' UTRs of plant r-protein mRNAs are rather short; for example, the 5' and 3' UTRs of tobacco *L2* are only 7 and 128 nt, respectively (Marty and Meyer, 1992). The petunia *L15* and *L27a* mRNAs have 248 and 176 nt 3' UTRs (Lee et al., 1999), the Arabidopsis *S15* mRNA has a 207 nt 3' UTR and an 83 nt 5' UTR (Sangwan et al., 1993), mRNAs transcribed from the two Arabidopsis *L23a* genes have 5' UTRs between 48 and 74 nt (McIntosh and Bonham-Smith, unpublished data; Chapter 4, this volume), and the 5' UTRs of maize *S6* mRNAs are only 69 nt (Williams et al., 2003). A distinctive feature of plant r-protein genes with characterized 3' UTRs is that they lack the canonical polyadenylation signals upstream of their poly $(A)^+$ tails that are present in animal mRNAs (Marty and Meyer, 1992; Sangwan et al., 1993; Lee et al., 1999; Giannino et al., 2000). Peach *S28* mRNAs have short (8-11 nt) 5' oligopyrimidine tracts (Giannino et al., 2000), as do the mRNAs encoding maize S6 (Williams et al., 2003). A preliminary report in an Arabidopsis genomic survey suggests that numerous r-protein genes from this species may show the 5' TOP trait (Barakat et al., 2001). *In vitro*, wheat germ ribosomes recognize and translate mammalian 5' TOP mRNAs with the same efficiency as quiescent cells, reflecting the polysomal association of the mRNA *in vivo*; the same results are obtained when rabbit reticulocyte ribosomes are used (Shama and Meyuhas, 1996). The effect of the 5' TOP sequence was abolished upon mutation of the motif, indicating specific recognition of 5' TOP mRNA by the plant ribosome that suggests a similar translational regulatory mechanism may be present in plants (Shama and Meyuhas, 1996). In a recent study of maize embryos, auxin was found to specifically stimulate both r-protein translation and phosphorylation of S6 (Beltrán-Peña et al., 2002). The specific recruitment of the 5' TOP-like mRNA for S6 into polysomes was stimulated by auxin treatment, increased by augmented S6 phosphorylation, and repressed by inhibited S6 phosphorylation, suggesting that plants have a similar r-protein translational control mechanism to that of animals (Beltrán-Peña et al., 2002).

Induction of S6 kinase activity and S6 phosphorylation was also shown to be correlated with an increase in r-protein mRNAs in polysomes in Arabidopsis cell cultures. As in animals, this induction and subsequent upregulation of r-protein mRNA translation was inhibited by a mammalian PI3 kinase inhibitor (Turck et al., 2004).

1.5.4.3. Post-translational regulation

Expression of multiple isoforms of each r-protein along with post-translational modification suggests a large degree of heterogeneity among plant ribosomes. Post-translational modifications have been documented for numerous r-proteins of the *E. coli* ribosome (Arnold and Reilly, 1999) and human 60S (Odintsova et al., 2003) and 40S (Yu et al., 2005) subunits, most commonly N-terminal loss of methionine, acetylation, and methylation. Recent analysis of Arabidopsis ribosomes from single tissue types, cell culture (Chang et al., 2005) and rosette leaf (Giavalisco et al., 2005), revealed a striking array of heterogeneity among ribosomes, with numerous r-proteins (~50% of identified r-proteins) represented by two or more gene family members and/or multiple 2D gel migrations (post-translational modifications). Ribosomal heterogeneity has also been identified in maize, where ribosomal incorporation and phosphorylation of acidic r-protein (P-protein) isoforms differed substantially between tissues (Szick-Miranda and Bailey-Serres, 2001).

Phosphorylation is, by far, the best characterized r-protein modification, and abiotic stresses and developmental stimuli often result in changes to the phosphorylation states of r-protein S6 and the acidic r-proteins. In general, translation is downregulated via reduction in ribosome biogenesis and/or translational efficiency (reflected in a decrease in polysomes and an increase in monosomes and half-mers) following a number of stresses, including heat (Fehling and Weidner, 1986; Nover et al., 1986; Scharf and Nover, 1987; Fehling and Weidner, 1988), cold (Bixby and Brown, 1975), and hypoxia (Bailey-Serres and Freeling, 1990; Fennoy and Bailey-Serres, 1995). Changes in r-protein composition, phosphorylation, and/or r-protein electrophoretic mobility of fractionated ribosomes following induction of heat- (Fehling and Weidner, 1988) or cold-tolerance (Bixby and Brown, 1975) suggest that r-protein modification is an important part of translational regulation in the plant cell. Heat stress in tomato culture (*Lycopersicon peruvianum*; Scharf and Nover, 1982) and maize embryos

(Beltrán-Peña et al., 2002), and hypoxia in maize seedling roots (Bailey-Serres and Freeling, 1990), have been reported to result in a decrease in phosphorylation of a 30-31 kDa r-protein of the 40S subunit, subsequently identified as S6. Release from heat stress was shown to result in S6 re-phosphorylation within two hours, with a subsequent increase in mitotic activity (Scharf and Nover, 1982). While phosphorylation of S6 decreases following stress phosphorylation of acidic r-proteins remains unchanged (Scharf and Nover, 1982; Scharf and Nover, 1987; Bailey-Serres and Freeling, 1990). Phosphorylation of r-protein S6 and the acidic r-proteins has also been found to fluctuate during development in maize embryos during a 24 h germination period (Pérez-Méndez et al., 1993). Auxin treatment also resulted in an increase in r-protein phosphorylation in germinating maize embryos (Pérez et al., 1990), specifically r-protein S6 (Beltrán-Peña et al., 2002).

R-protein S6 is encoded by two genes in maize, *RPS6A* and *B*, but post-translational modification, phosphorylation at any of six C-terminal serine and threonine residues (Williams et al., 2003), adds a great deal of complexity to the regulation of expression of the two genes. Nine different S6 isoforms in maize seedlings have been identified, including non-, mono-, di-, tri-, tetra-, and pentaphosphorylated forms, all found in ribosomes of untreated seedling root tips (Williams et al., 2003). Stress treatments resulted in changes to the relative abundance of each isoform; heat and anoxia resulted in an increase in non-, mono-, and diphosphorylated S6 and a decrease in levels of tri-, tetra-, and pentaphosphorylated S6. In contrast, cold stress resulted in the accumulation of the hyperphosphorylated forms of S6 (Williams et al., 2003). Under non-stress and anoxia conditions, a Ser/Thr phosphatase is responsible for depletion of the hyperphosphorylated forms of S6, while inhibition of the phosphatase under cold stress conditions results in the accumulation of hyperphosphorylated S6 (Williams et al., 2003).

How S6 phosphorylation is regulated in plants presents an interesting question. Transcript levels of Arabidopsis S6 kinase are induced under auxin treatment and both cold and salinity stresses (Mizoguchi et al., 1995). An Arabidopsis homolog of human $p70^{s6k}$, AtS6k2, not only demonstrates S6-specific kinase activity in both human and plant cells, but is also inhibited by heat stress, concurrent with S6 dephosphorylation

(Turck et al., 1998). Although AtS6k2, like $p70^{s6k}$, is regulated by phosphorylation of key Ser and Thr residues, unlike $p70^{s6k}$, it is not sensitive to rapamycin (Turck et al., 1998; Turck et al., 2004). Arabidopsis has a single TOR kinase, AtTOR, expressed in embryos and meristems but not in differentiated cells, which is essential for both embryo and endosperm development (Menand et al., 2002). It is not known how the AtTOR signaling pathway functions in Arabidopsis, but it is likely that the S6 kinase pathway in plants does not involve AtTOR, given the rapamycin insensitivity (Turck et al., 1998) of AtS6k2 (Menand et al., 2002). The recovery of S6 hyperphosphorylation following anoxic stress in maize (Williams et al., 2003), and the stimulation of AtS6k activity and S6 phosphorylation in Arabidopsis cultures stimulated by fresh medium (Turck et al., 2004), are both inhibited by an inhibitor of mammalian PI3 kinase, LY40092. Despite the lack of a plant homolog of mammalian PI3K, a related plant lipid kinase must be part of the signaling pathway that activates S6k, stimulating S6 phosphorylation (Turck et al., 2004). The dependence of S6k activation on phytohormones (auxin, cytokinin), and the differential regulation of the two AtS6k isoforms in response to growth stimuli (Turck et al., 2004) indicate that there are many unique aspects to S6 post-translational regulation and r-protein translational regulation in plants.

1.6. Ribosomal protein L23a

RPL23a is a member of one of the 32 r-protein families conserved across all domains of life (Lecompte et al., 2002), with an evolutionary history of over two billion years since the last universal common ancestor (Vishwanath et al., 2004). Eukaryotic r-protein L23a was first identified (as L23') by 2D electrophoresis of the 60S subunit of the rat liver ribosome (Tsurugi et al., 1977). L23a and its yeast equivalent, L25, have subsequently been identified and sequenced initially in yeast (Bollen et al., 1982), rat (Suzuki and Wool, 1993), and humans (Fan et al., 1997), and now in many other eukaryotes. L23a and L25 are the eukaryotic counterparts of the *E. coli* r-protein, L23. An *L23a* cDNA isolated from *A. thaliana*, now designated *AtRPL23aA*, was cloned and sequenced by Bonham-Smith (1997, GenBank accession no. AF034694) and the full-length gene for *RPL23aA* has also been cloned and identified (At2g39460). Following the sequencing efforts of the Arabidopsis Genome Initiative (2000), a genomic clone for another *AtRPL23a* gene, now designated *AtRPL23aB*, was identified (At3g55280). A

complete survey of the Arabidopsis r-protein genes shows that *AtRPL23aA* and *B* are the only genes encoding Arabidopsis L23a; as indicated by the presence of cDNAs, both copies are expressed (Barakat et al., 2001).

Prokaryotic (L23) and eukaryotic (L25, L23a) homologues are members of the highly conserved L23/L25 (or L23p; Lecompte et al., 2002) r-protein family. The L23/L25 family members directly bind to a conserved site in domain III of 23S or 23S-like rRNA (Buisson and Reboud, 1982; Vester and Garrett, 1984; El-Baradi et al., 1984, 1985, 1987; Jeeninga et al., 1996), playing a role in the formation of the large ribosomal subunit. The *S. cerevisiae* L25 domain responsible for 26S rRNA binding has been identified (Rutgers et al., 1991), and there is a corresponding conserved domain in other L23/L25 family members. The secondary and tertiary structure of the L23/L25 binding site is highly conserved between 23S/23S-like rRNA of different species, and contact points between r-protein and rRNA have been well-characterized (Egebjerg et al., 1991; Kooi et al., 1993; van Beekvelt et al., 2000). The L23 binding site on 23S rRNA in *E. coli* was predicted to be close to the peptidyl transferase center of the ribosome, but is not directly at the active site (Weitzmann and Cooperman, 1990). Adjacent to r-proteins L29 and L39e, *H. marismortui* L23 is located near the ribosome's polypeptide exit tunnel, suggesting a possible role in protein secretion during contact with the translocon of the ER (Nissen et al., 2000).

Transport of eukaryotic L23/L25 r-proteins to the nucleus and nucleolus following translation is a key step in ribosome biogenesis. Any one of 4 different importin β-like transport receptors (importin β, transportin, RanBP5, and RanBP7) was reported to import L23a from canine pancreas ribosomes (Jakel and Gorlich, 1998). *S. cerevisiae* L25 has an N-terminal nuclear localization signal (NLS) and may play a key role in assembly of the large ribosomal subunit as one of the specific binders of 26S rRNA (Schaap et al., 1991). *S. cerevisiae* L25 has also been determined to play a role in the processing of the 27S pre-rRNA (van Beekvelt et al., 2001).

Perhaps the best-characterized function of L23/L25 homologues is their role in ribosome docking and the secretory pathway. *H. marismortui* L23 has been identified as one of the r-proteins surrounding the exit tunnel (Ban et al., 2000; Nissen et al., 2000) that interacts with the prokaryotic chaperone Trigger Factor (Kramer et al., 2002; Blaha

et al., 2003). Components of the signal recognition particle in both prokaryotes (Gu et al., 2003; Kramer et al., 2002) and eukaryotes (Pool et al., 2002) have also been identified as L23/L25 interactors. The crystal structure of eukaryotic SRP docked to the ribosome confirms that a subunit of SRP contacts L23a (Halic et al., 2004). When the ribosome is docked at the translocon of the ER membrane during co-translational translocation, the ribosome is positioned with the peptide exit tunnel directly over top of the translocon pore, with connections involving L23a/L25 (Beckmann et al., 2001; Morgan et al., 2002).

1.7. Objectives

The long-term objective of this research is to understand how coordinate regulation of plant r-protein genes is achieved during ribosome biogenesis. In order to further this goal, the immediate objective of this thesis was to determine regulatory mechanisms controlling the expression of the members of a single plant r-protein gene family. Arabidopsis was chosen as a research subject due to its recognition as a model plant species for molecular research. R-protein L23a was chosen for study because of its homology to a family of primary binding r-proteins with known roles in assembly of the large subunit. The objectives of this research were to: 1) establish AtRPL23a as a functional r-protein, 2) characterize and compare the expression patterns of both *AtRPL23a* genes, and 3) delineate and compare *cis*-elements involved in the regulation of *AtRPL23aA* and *B*.

CHAPTER 2. ESTABLISHMENT OF *ARABIDOPSIS THALIANA* RIBOSOMAL PROTEIN RPL23AA AS A FUNCTIONAL HOMOLOGUE OF *SACCHAROMYCES CEREVISIAE* RIBOSOMAL PROTEIN L25

Arabidopsis thaliana ribosomal protein (r-protein) RPL23aA shows 54% amino acid sequence identity to the *Saccharomyces cerevisiae* equivalent r-protein, L25. *AtRPL23aA* also shows high amino acid sequence identity to members of the L23/L25 r-protein family in other species. R-protein L25 in *S. cerevisiae* has been identified as a primary rRNA binding protein that directly binds to a specific site on yeast 26S rRNA. It is translocated to the nucleolus where it binds to 26S rRNA and plays an important role in ribosome assembly. Transformation of a *S. cerevisiae l25* mutant with the *AtRPL23aA* cDNA rescued the mutant phenotype. This work establishes the first isolated AtRPL23a gene as a functional equivalent of yeast L25 via rescue of a yeast *l25* mutant strain. This establishes Arabidopsis L23a as the first demonstrated functional plant member of the conserved L23/L25 r-protein gene family, with all of the attendant functions demonstrated by the *S. cerevisiae* member of this family, L25. By rescuing the *l25* strain, AtRPL23aA must be able to function in the capacity of L25 in yeast ribosomes. The AtL23aA nuclear localization signal must also have been recognized by the yeast cell in order for transport back into the nucleus, and ribosome biogenesis.

2.1. Introduction

Ribosomes are the large ribonucleoprotein complexes responsible for polypeptide synthesis in all living cells. Eukaryotic ribosomes consist of four ribosomal RNAs (rRNA) and over 70 ribosomal proteins (r-proteins). The ribosomes of *Saccharomyces cerevisiae* contain approximately 75 r-proteins (Warner, 1989), rat ribosomes contain approximately 80 r-proteins (Wool et al., 1995), and plant

cytoplasmic ribosomes contain between 75 and 92 r-proteins, depending upon the species (Bailey-Serres, 1998). Given the universal requirement for ribosome function, it is not surprising that the ribosomes of all organisms, both prokaryotic and eukaryotic, are highly similar. Accordingly, some ribosomal constituents have structural features that are highly conserved between species and across kingdoms. Many aspects of the secondary structure of the large subunit high molecular weight rRNA (23S rRNA in *Escherichia coli*, 23S-like rRNA in eukaryotes, i.e. 26S, 28S) are highly conserved, despite differences in rRNA size between species (reviewed in Raué et al., 1988; Schnare et al., 1996).

Rigorous r-protein removal experiments using proteinase K, SDS, and phenol extraction treatments, among others, have suggested that the peptidyl transferase function of the ribosome is carried out primarily by the large subunit rRNA (Noller et al., 1992; Noller, 1993). Crystallographic studies have provided structural evidence that the large subunit rRNA functions as the peptidyl transferase, confirming the ribosome as a functional ribozyme (Nissen et al., 2000). In addition, the 23S (or 23S-like) rRNA, as well as the other large subunit (5S, 5.8S) and small subunit rRNAs (16S in *E. coli*, 18S in eukaryotes) have crucial roles to play in ribosome structure. Together with the rRNAs, the r-proteins have an important, yet poorly understood, role in ribosome structure and function. Some r-proteins are known to be necessary for peptidyl transferase activity. Prokaryotic r-protein L16, which has no peptidyl transferase function on its own, can restore this activity to nonfunctional 50S ribosomal subunits stripped of r-proteins (Moore et al., 1975). A useful approach to elucidate the function of r-proteins within the ribosome is to first study those r-proteins that interact directly with the rRNAs. One such r-protein, L25 in *S. cerevisiae*, is a primary rRNA binding protein that associates directly with 26S rRNA early in ribosome large subunit assembly (El-Baradi et al., 1984).

Assembly of the eukaryotic ribosome is a complex process requiring the coordinated expression and transport of all the constituent rRNAs and r-proteins. R-protein genes are transcribed in the nucleus and the resulting messenger RNAs (mRNAs) are transported to the cytosol for translation, with the majority of the resulting r-proteins being subsequently transported to the nucleolus for ribosome subunit

assembly to occur. Nuclear import of r-proteins is therefore an essential part of ribosome subunit assembly, and nuclear localization signals (NLS) in some r-proteins are required for this process. *S. cerevisiae* r-protein L25 contains NLS sequences (Schaap et al., 1991). Once translocated to the nucleus, L25 is thought to play a key role in large subunit assembly, as binding of L25 to 26S rRNA may be required in order for further r-protein accumulation to occur (El-Baradi et al., 1985).

S. cerevisiae L25 is part of a conserved group of r-proteins, the L23/L25 family, which contains members from both prokaryotic and eukaryotic species. The 23S rRNA binding site of the L25 homologue, L23 in *E. coli,* is perhaps the best characterized of the L23/L25 binding sites; it is on domain III of the 23S rRNA near the peptidyl transferase center of the ribosome (Vester and Garrett, 1984; Weitzmann and Cooperman, 1990; Thiede et al., 1998). Structural studies of *Haloarcula marismortui* ribosomes specifically place L23 near the ~100 Å long polypeptide exit tunnel, the inner surface of which consists largely of 23S rRNA domains I to V (Nissen et al., 2000). L23 interacts with domain III of the 23S rRNA, and is adjacent to r-proteins L29 and L39e (Nissen et al., 2000). The location of r-proteins near the opening of the exit tunnel suggests a possible role in protein secretion, since such r-proteins are in a position to associate with the translocon of the endoplasmic reticulum during protein secretion (Nissen et al., 2000). The 26S rRNA binding site for *S. cerevisiae* L25 has likewise been well defined using ribonuclease protection and mutagenesis experiments (van Beekvelt et al., 2000). Sequence comparison studies provide evidence that the L23/L25 r-proteins and their corresponding binding sites on large subunit rRNA have been highly conserved over evolutionary time. In addition to the structural conservation seen in sequence comparisons, functional studies have demonstrated the ability of bacterial, yeast, and mammalian L23/L25 r-proteins to bind the large subunit rRNA of other species (El-Baradi et al., 1985, 1987; Jeeninga et al., 1996). Rat r-protein L23a is able to rescue *S. cerevisiae l25* mutants, and in doing so compete with native L25 for assembly into yeast large ribosomal subunits (Jeeninga et al., 1996).

In this report, we describe the first complementation of a yeast *l25* mutant with a plant r-protein, *RPL23a*A from *Arabidopsis thaliana*. *AtRPL23a*A is the first plant r-protein to be placed experimentally in the L23/L25 family of conserved r-proteins,

demonstrating functional conservation of the L23/L25 r-proteins across all major kingdoms.

2.2. Materials and Methods

2.2.1. Strains and selection

Escherichia coli strain DH5α was used as a bacterial host for cloning/plasmid construction. *Saccharomyces cerevisiae* strain YCR61 [*MAT***a**, *ade1-100, leu2-3,112, his4-519 ura3-52, GAL::rpL25* (*URA3*)] (formerly BWG1-7A/42; Rutgers et al., 1990) was kindly provided by Dr. Jaap Venema (Vrije Universiteit, Amsterdam). YCR61 contains the r-protein gene *L25* behind the *GAL1-10* upstream activating sequence (Rutgers et al., 1990) and conditionally expresses r-protein L25 when grown on synthetic galactose (SGal) medium without uracil. Transformed YCR61 lines were grown on glucose-based (synthetic dextrose, SD) medium; both YPD rich medium and SD minus tryptophan (SD -trp) selection medium were used. All yeast media was prepared according to Sherman et al. (1983) and all incubation was carried out at 30^{o}C. Galactose-based medium was prepared by substituting D-galactose (Sigma) for D-glucose (BDH).

2.2.2. **A. thaliana** *cDNA and amino acid sequence alignments*

The *A. thaliana* EST for *AtRPL23aA* (accession #R84194) was obtained from the Arabidopsis Biological Resource Center (Ohio State) and sequenced via the dideoxy chain termination procedure (Sanger et al., 1977). The Arabidopsis cDNA amino acid sequence (*AtRPL23aA*; accession no. AAB87692, translated from cDNA accession no. AF034694) was compared to *AtRPL23aB* (accession no. CAB75762, translation of chromosome III *RPL23a* genomic sequence accession no. AL132954) and homologues from *S. cerevisiae* (P04456, CAA99146 translated from gene coding sequence Z74869), *Rattus rattus* (CAA46336, translated from cDNA X65228), *Homo sapiens* (AAC51934, translated from genomic sequence NM_020217), *Nicotiana tabacum* (AAA53296, translated from cDNA L18908), *E. coli* (CAA26462, translated from genomic sequence X02613), and *Methanococcus jannaschii* (AAB98163, translated from genomic sequence L77117/U67474) using ClustalW alignment software (Thompson et al., 1994; http://www.ebi.ac.uk/clustalw/).

2.2.3. Plasmids

The yeast shuttle vector pSCW231 (Watkins et al., 1993) was kindly provided by Dr. Wei Xiao (University of Saskatchewan). The *AtRPL23aA*-pSCW31 plasmid was constructed by ligating the *AtRPL23aA* cDNA into pSCW231, behind the alcohol dehydrogenase constitutive (*ADC1,* also called *ADH*) promoter, at *Eco*RI and *Bam*HI. pSCW231 carries the *TRP1* selection marker.

2.2.4. Transformation and confirmation

E. coli (DH5α strain) transformations were carried out via electroporation (BTX ECM399 electroporator). YCR61 was transformed with the *AtRPL23aA*-pSCW231 construct using a LiAC/ssDNA/PEG protocol (Agatep et al., 1998). YCR61 transformants were selected first by growth on YPD (glucose) medium followed by SD -trp medium to select for the presence of the pSCW231 (*TRP1*) construct.

Expression of the *AtRPL23aA*-pSCW231 construct in YCR61 transformants was confirmed using reverse transcription (RT) PCR. RNA was isolated from untransformed YCR61 and YCR61 transformants using a QIAGEN RNeasy kit, according to the manufacturer's Yeast II protocol (enzymatic lysis procedure). RT-PCR using ~2 μg RNA from transformed and untransformed YCR61 was carried out using a QIAGEN OneStep RT-PCR kit. Primers used in the RT-PCR amplification of the *AtRPL23aA* open reading frame (465 bp) are as follows : $^{5'}$GCGGGATCCATGTCTCCGGCTAAAGTTGAT$^{3'}$ and $^{5'}$GCGGGTACCTTAGATGATGCCGATCTTGTT$^{3'}$. Reverse transcription for 30 min at 50^{o}C (and subsequent heat inactivation of reverse transcriptase for 15 min at 95^{o}C) was follwed by 35 cycles of PCR (30 s at 94^{o}C, 30 s at 55^{o}C, 1 min at 72^{o}C).

2.2.5. Growth curves and rates

From overnight cultures of untransformed and transformed YCR61 in YPGal, 2.43×10^7 cells were added to 100 mL SGal -ura or 100 mL SD -trp media. OD_{600} readings (Beckman DU 7000 series spectrophotometer) were taken of 1 mL aliquots of culture at regular intervals over a 56 hour time course (readings every 2 hours from 8-20 hours, every 4 hours from 20-36 hours, and at 44 and 56 hours). Growth rates were calculated using slope readings from the resulting curves (Microsoft Excel). YCR61 was grown in 100 mL SD -trp medium (2.43×10^7 cells from YPGal medium were used to

start the culture) for 24 hours followed by centrifugation, then resuspended in 100 mL SGal -ura and incubated for a further 40 hours in order to determine recovery. OD_{600} readings for YCR61 recovery growth were taken using the same time points as for the other growth curves.

2.3. Results

2.3.1. AtRPL23aA shows high sequence identity to other L23/L25 r-proteins

Complete sequencing of an Arabidopsis EST confirmed its identity as *RPL23a*. Searches of GenBank showed the Arabidopsis genome to contain at least two *RPL23a* genes, on chromosomes II and III. The sequences of both *AtRPL23a* genes (present in BACs) are available through GenBank (accession nos. AC004218, AE002093; clones At2g39460 and At3g55280) in addition to the *AtRPL23a* cDNA sequence. A comparison of the *AtRPL23a* cDNA sequence to both *AtRPL23a* genes (alignment not shown) indicated that the cDNA was a product of the chromosome II copy of *AtRPL23a*.

Alignments of AtRPL23aA with L23/L25 amino acid sequences from other eukaryotes showed a high level of primary sequence conservation (Figure 2.1; Table 2.1). As expected, AtRPL23aA showed closest overall identity with AtRPL23aB, followed by *N. tabacum* L25, rat L23a, human L23A, and a relatively high level of identity (54%) with *S. cerevisiae* L25 (Table 2.1). A comparison of identified functional regions in the L23/L25 r-proteins, i.e. the NLS and C-terminal binding domain sequences, again showed a high level of conservation within these regions (Table 2.1). Functional regions were determined by comparing the known NLS (Schaap et al., 1991) and rRNA-binding domains (Rutgers et al., 1991; Kooi et al., 1994) from *S. cerevisiae* L25 with the corresponding sequences in the other L23/L25 r-proteins. AtRPL23aA showed 51% overall identity to the *S. cerevisiae* L25 NLS (amino acids 1-41; Schaap et al., 1991), with 60% identity to the NLS sub-region known to enhance efficiency of nuclear accumulation (amino acids 1-11) and 57% and 45% identity to the portions of the NLS required for NLS activity (amino acids 11-17 and 18-28 respectively; Schaap et al., 1991). AtRPL23aA also showed 60% identity to the region of L25 identified by Rutgers et al. (1991) as that involved in specific binding to yeast 26S rRNA (amino acids 62-142). Fifty-seven percent identity was seen between AtRPL23aA and L25 in

Sequence for comparison	Overall identity (%)	NLS identity (%)	C-terminal identity (%)	GenBank accession #
A. thaliana RPL23aB	94	89	99	CAB75762
N. tabacum L25	84	72	98	AAA53296
H. sapiens L23A	71	56	81	AAC51934
R. rattus L23a	71	56	81	CAA46336
S. cerevisiae L25	54	51	60	P04456
M. jannaschii L23	47	-	48	AAB98163
E. coli L23	27	-	30	CAA26462

Table 2.1. Amino acid sequence comparison between AtRPL23aA and L23/L25 r-protein sequences from other species. NLS identity refers to identity between regions corresponding to amino acids 1-41 from *S. cerevisiae* L25 (Schaap et al., 1991). C-terminal rRNA binding domain identity was determined using r-protein regions corresponding to *S. cerevisiae* L25 amino acids 62-142 (Rutgers et al., 1991).

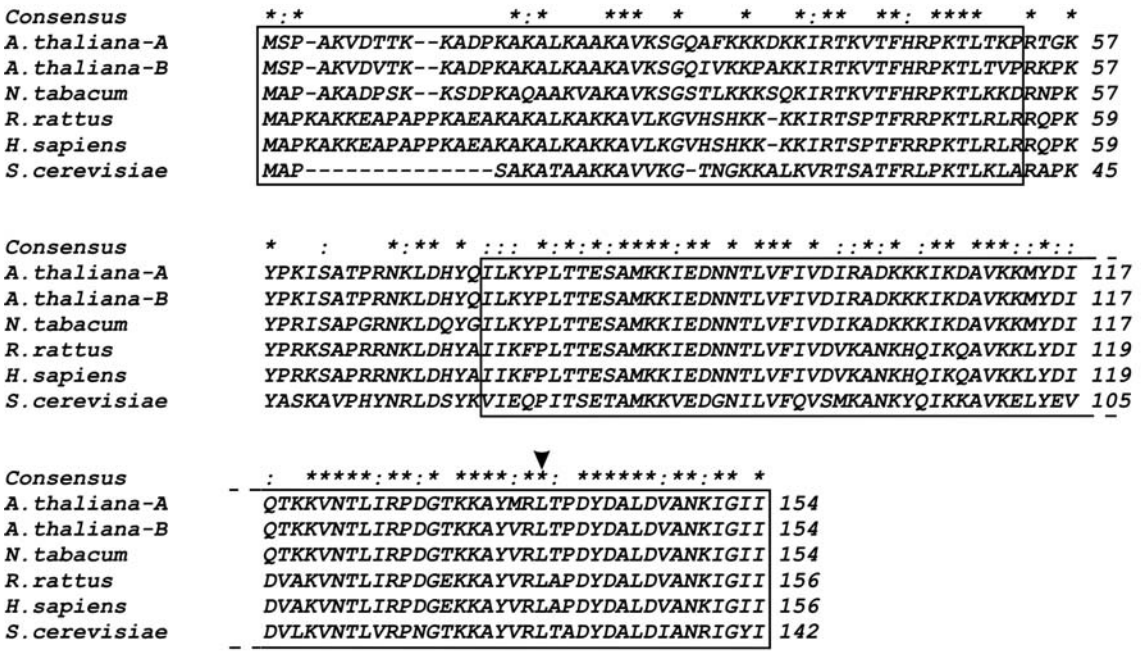


Figure 2.1. Alignment of eukaryotic L23a/L25 r-proteins (ClustalW). Boxed regions indicate NLS and C-terminal rRNA binding domains as identified in *Saccharomyces cerevisiae* L25 (NLS, amino acid residues 1-41, Schaap et al., 1991; rRNA binding domain, residues 62-142, Rutgers et al., 1991; residue numbering according to *S. cerevisiae* L25 sequence, amino acids 1-142). NLS sub-regions known to enhance the efficiency of nuclear accumulation (residues 1-11) and those required for NLS activity (residues 11-17 and 18-28; Schaap et al., 1991) in *S. cerevisiae* L25 are included in the boxed NLS region. C-terminal sub-regions required for specific binding of 26S rRNA (residues 62-126) and enhancing binding efficiency (residues 127-142; Rutgers et al., 1991) in *S. cerevisiae* L25 are included in the boxed C-terminal domain. The key leucine residue (leu 126) required for L25 to 26S rRNA binding in *S. cerevisiae* (Rutgers et al., 1991) and its counterparts in the other L23/L25 sequences is indicated by an arrow. For GenBank accession numbers, see Table 1. *A. thaliana*-A, RPL23aA; *A. thaliana*-B, RPL23aB; *, identical amino acids in all sequences; :, conserved substitutions (Thompson et al., 1994).

the region required for specific binding of 26S rRNA (amino acids 62-126) and 80% identity was conserved in the region responsible for enhancing binding efficiency (L25 amino acids 127-142); notably, leucine 126 of L25, known to be crucial for rRNA binding (Rutgers et al., 1991), was found to be conserved among all L23/L25 amino acid sequences compared, including that of AtRPL23aA (Figure 2.1).

The L23/L25 r-proteins of the eukaryotic species compared all showed less conservation of sequence in the NLS region (44% identity) than throughout the C-terminal rRNA binding domain (58% identity; Table 2.1). In a comparison of overall sequence, NLS, and rRNA binding domain sequences, higher levels of identity are shown between AtRPL23aA and the L23/L25 r-proteins of other multicellular eukaryotes than between AtRPL23aA and *S. cerevisiae* L25. Interestingly, a sequence comparison between AtRPL23aA and *E. coli* L23 showed less overall identity (27%) than a comparison between AtRPL23aA and the corresponding L23 r-protein from the archaeon, *M. jannascii* (47%; Table 2.1).

2.3.2. AtRPL23aA complements a yeast l25 mutant

In order to determine functional complementarity, the *AtRPL23aA* cDNA, including 5' and 3' untranslated sequence, was placed under the control of the constitutive *ADH1* promoter in the shuttle vector pSCW231 (*TRP1*); this construct was used to transform YCR61, a *S. cerevisiae* strain that expresses a single copy of *L25* only when grown on a galactose medium (Rutgers et al., 1990). Transformants were selected first on YPD medium and then transferred to SD -trp plates (Figure 2.2). Untransformed YCR61 did not grow on YPD medium (data not shown). Five transformant colonies were successfully grown on SD -trp medium and *AtRPL23aA* expression was confirmed using RT-PCR to amplify the 465 base pair *AtRPL23aA* open reading frame from total RNA of transformant YCR61 (Figure 2.3).

2.3.3. Growth of YCR61 transformants

As expected, transformant YCR61 lines were able to grow on both galactose- and glucose- based media, however, they grew more slowly on SGal medium (when both *AtRPL23a*A and L25 were expressed) than untransformed YCR61 (Figure 2.2). Growth characteristics of transformed YCR61 were determined and compared to that of untransformed YCR61 in both SGal -ura and SD -trp media (Figure 2.4A, B). At the

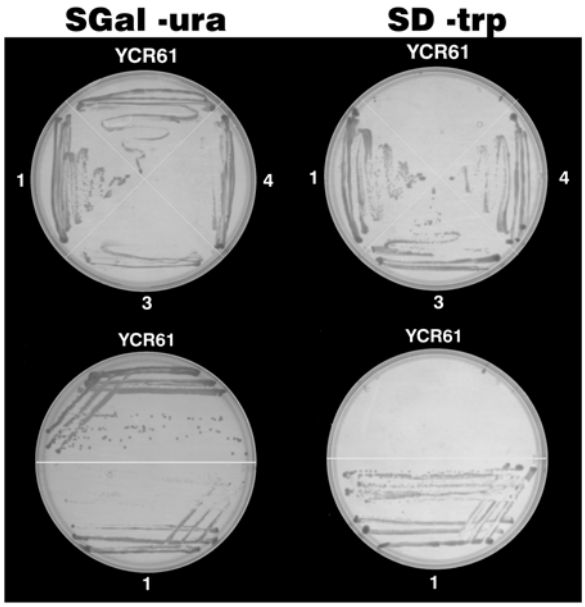


Figure 2.2. Growth characteristics of YCR61 when transformed with *AtRPL23aA* in the shuttle vector pSCW231. Growth of untransformed YCR61 and transformants YCR61-1 (1), YCR61-3 (3), and YCR61-4 (4) is shown on both SGal (non-selective) and SD (selective) medium. Plates were incubated for 3 days at 30^{o}C.

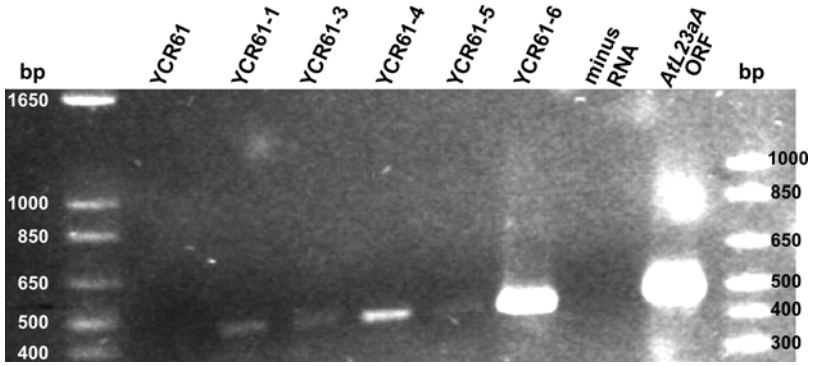


Figure 2.3. RT-PCR of untransformed YCR61 and transformed YCR61 lines (YCR61-1, YCR61-3, YCR61-4, YCR61-5, YCR61-6) after amplification of the *AtRPL23aA* open reading frame (465 bp). DNA size markers are indicated in bp to the right and left of the gel.

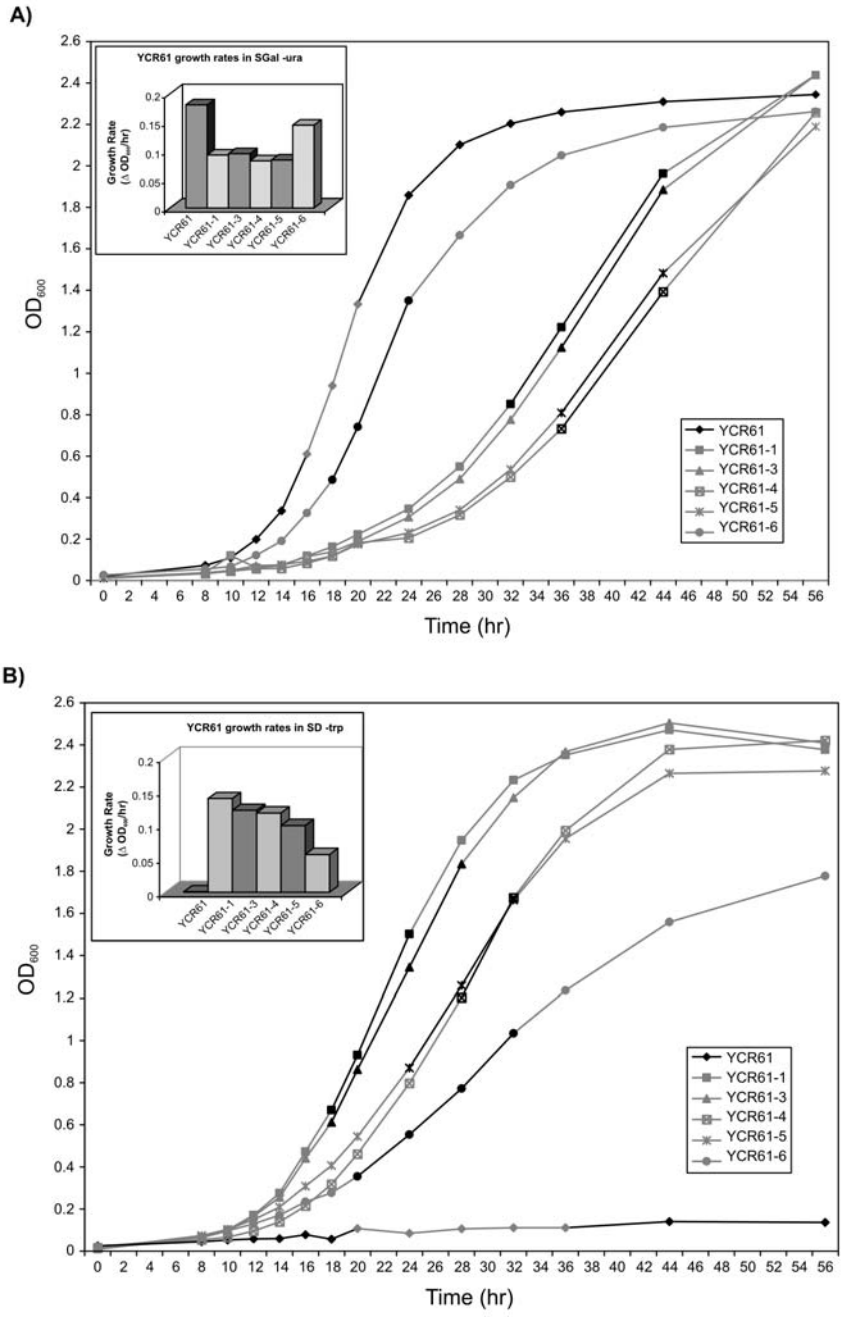


Figure 2.4. Growth curves for untransformed and transformant YCR61. Curves represent an average of two independent experiments. A) YCR61 growth curve in SGal –ura medium; inset: average growth rate calculated using the slope of the line in the region indicated in gray (YCR61) or black (YCR61 transformants) on the line graph. B) YCR61 growth curve in SD -trp medium; inset: average growth rate calculated using the slope of the line in the region indicated in gray (YCR61) or black (YCR61 transformants) on the line graph.

completion of the 56 hour time course, both YCR61 (average growth rate of $1.807x10^{-1}$ ΔOD_{600}/hr) and YCR61-6 (average growth rate of $1.453x10^{-1}$ ΔOD_{600}/hr) grown in SGal -ura had reached stationary phase, while the four remaining transformant YCR61 lines (average growth rates of $8.257x10^{-2}$ – $9.515x10^{-2}$ ΔOD_{600}/hr) were completing log phase growth. As expected, untransformed YCR61 showed virtually no growth (average growth rate of $8.620x10^{-4}$ ΔOD_{600}/hr) in the SD -trp medium (Figure 2.4B). YCR61-1, at $1.393x10^{-1}$ ΔOD_{600}/hr and YCR61-3 at $1.221x10^{-1}\Delta OD_{600}$/hr had the highest average growth rates in SD -trp, and all transformant lines except YCR61-6 were into stationary growth by the end of the time course (Figure 2.4B). YCR61-6, which had the fastest growth rate of the transformant lines in SGal -ura, showed the slowest average rate of growth ($5.632x10^{-2}$ ΔOD_{600}/hr) in SD -trp. The recovery curve for untransformed YCR61 showed an increase in growth after the cells were transferred from the growth-inhibiting SD -trp medium to SGal -ura medium (data not shown). Morphology of all untransformed and transformant YCR61 lines in both medium types was monitored during the time course. All YCR61 transformants except YCR61-5 appeared normal in morphology, however, many YCR61-5 cells exhibited an elongated phenotype with abnormal division (unlike normal budding). The YCR61-5 phenotype has been previously documented by Schaap et al. (1991) where it was observed in transformed YCR61 expressing L25 minus the NLS (amino acids 2 to 41).

2.4. Discussion

This work introduces the first plant r-protein to be experimentally confirmed as a member of the L23/L25 r-protein family. Amino acid sequence alignments demonstrate that AtRPL23aA shows structural similarities to its L23/L25 homologues. As well as AtRPL23aA showing high overall sequence identity to other eukaryotic L23/L25 r-proteins, it has also retained key sequence motifs in the NLS and rRNA binding domains (Figure 2.1; Table 2.1). In addition to being a structural equivalent to previously identified L23/L25 r-proteins, successful complementation of the *S. cerevisiae l25* mutant, YCR61, by AtRPL23aA clearly demonstrated the functional equivalence of L25 and *AtRPL23a*A. Complementation of the YCR61 mutant by AtRPL23aA not only infers the ability of the AtRPL23aA NLS to be recognized by the *S. cerevisiae* transport machinery, but also demonstrated the ability of AtRPL23aA to

bind yeast 26S rRNA, fulfilling a critical role as a primary, direct binder of rRNA. By demonstrating the functional equivalence of L25 and AtRPL23aA, we have established a role for ribosomal protein RPL23aA in the Arabidopsis ribosome.

The ability of *AtRPL23a*A to function as a primary binder of domain III of 26S rRNA reflects the high degree of conservation between the secondary structures of *S. cerevisiae* 26S rRNA and Arabidopsis 25S rRNA, as illustrated in Figure 2.5. The tertiary interaction in domain III of the 26S rRNA shown to be required for L25 binding (Kooi et al., 1993) is also conserved in Arabidopsis 25S rRNA. These data add to previous findings establishing the ability of L23/L25 r-proteins to bind heterologous large subunit rRNAs (El-Baradi et al., 1985; El-Baradi et al., 1987; Jeeninga et al., 1996).

Expressing *AtRPL23aA* under the control of a constitutive promoter did not allow r-protein expression to be regulated in a growth- or environment-dependent manner, as would be the case in wild type yeast. We suggest that the slower growth of theYCR61 transformants on galactose medium, when compared to untransformed YCR61, may have been due to competition between the endogenous *S. cerevisiae* L25 and AtRPL23aA when both were being expressed. Although *AtRPL23aA* was able to complement the *l25* mutant, the two r-proteins are not identical; differences in primary protein structure may reflect different binding abilities to rRNA or other r-proteins, even if primary rRNA binding ability is conserved. In the case of L23/L25 r-proteins, structural differences between r-proteins may affect the ability of the r-protein to function in the established roles of rRNA processing (van Beekvelt et al., 2001), chaperone docking (Kramer et al., 2002; Pool et al., 2002; Blaha et al., 2003; Gu et al., 2003) and docking the ribosome to the ER (Beckmann et al., 2001; Morgan et al., 2002).

The slower growth of YCR61 transformants may also have been due to the inability of the yeast nucleocytoplasmic transport machinery to fully recognize the AtRPL23aA NLS; although key NLS motifs are conserved between L25 and AtRPL23aA, overall NLS identity between the two r-proteins is only 51% (Table 2.1).

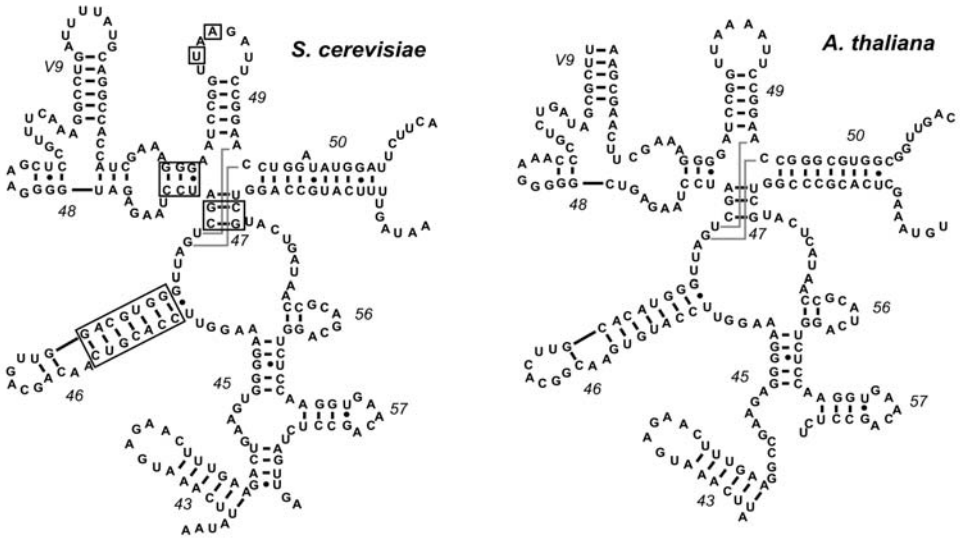


Figure 2.5. Predicted secondary structure of Domain III of *S. cerevisiae* 26S rRNA (adapted from van Beekvelt et al., 2000) and *A. thaliana* 25S rRNA (accession no. X52320). Tertiary interactions in domain III, required for L25 to 26S rRNA binding in *S. cerevisiae* (Kooi et al., 1993), are shown with gray lines. Regions of 26S rRNA domain III involved in L25 binding as determined by 26S mutagenesis and *in vitro* L25 binding studies (van Beekvelt et al., 2000) are boxed. Numbering of the helices is according to Raué et al. (1988).

Previous studies have shown that transformant yeast expressing copies of L25 lacking a functional NLS grew more slowly and showed the same abnormal phenotype as we report in YCR61-5 cells; it was suggested that the abnormal phenotype was due to a lack of translational efficiency in the translation of cell division proteins, resulting in inefficient cell division (Schaap et al., 1991). Native r-proteins are known to be preferentially assembled into ribosomes when they are competing with homologous r-proteins from other species, and slow growth phenotypes can result when both native and heterologous r-proteins are present *in vivo* (Fleming et al., 1989; Dick and Trumpower, 1998).

In this work we used the entire *AtRPL23aA* cDNA, including the 5' and 3' UTRs, in the expression vector used to complement the *l25* mutation. Previous complementation of the YCR61 mutant with RPL23a from rat utilized only the open reading frame of the rat r-protein under the control of the *S. cerevisiae L25* promoter and terminator sequences (Jeeninga et al., 1996). Complementation of yeast *qsr1* (*rpL10*) mutants with human or corn homologous r-protein cDNAs also utilized only the open reading frames of the heterologous r-proteins (Dick and Trumpower, 1998). Behind *ADH* constitutive promoters, corn or human *QM* (*rpL10*) sequences, containing their corresponding 5' and 3' untranslated regions, failed to complement yeast *qsr1* mutants; constructs containing the human or corn *QM* open reading frame flanked by yeast *QSR1* 5' and 3' untranslated regions behind the yeast *QSR1* promoter did, however, successfully complement the *qsr1* mutation (Dick and Trumpower, 1998). The previous rat, human, and corn r-protein complementation experiments suggested that successful complementation of yeast mutants required native yeast regulatory sequences controlling expression of the heterologous r-protein transcripts. In contrast, our data demonstrated successful complementation of a yeast *l25* mutant when 5’ and 3’ untranslated regions of the *AtRPL23aA* cDNA were included in the transformation construct. RT-PCR confirmed the presence and expression of the *AtRPL23aA*-pSCW231 construct in transformant YCR61. That transformants were able to grow on SD –trp medium indicates successful translation of the *AtRPL23aA* transcripts occurred. These data suggest that the 5’ and 3’ untranslated regions of the *AtRPL23aA* cDNA did not interfere

with efficient expression of the r-protein cDNA, and that the *AtRPL23aA* transcripts must have been recognized by the yeast translational machinery.

Chapter 3. THE TWO RIBOSOMAL PROTEIN *L23A* GENES ARE DIFFERENTIALLY TRANSCRIBED IN *ARABIDOPSIS THALIANA*

The two AtRPL23a isoforms, RPL23aA and RPL23aB, are 94% identical at the amino acid level yet *RPL23aA* and *B* share only ~40-50% primary sequence identity within the 5' regulatory regions. While the *RPL23aA* and *B* 5' regulatory regions share many similar predicted motifs, the arrangement and number of these motifs differs between the two genes. Differences in regulation between *RPL23aA* and *B* have been investigated via reverse transcription-PCR (RT-PCR) expression profiles. Overall, transcript abundance for *RPL23aA* and *B* varied slightly in specific tissues and under some abiotic stresses. The highest transcript abundance for both *RPL23a* genes was detected in mitotically active tissues such as bud, flower and elongating carpel, as well as in root and stem while the lowest transcript levels were found in mature leaf and bract. Hormone-treated seedlings showed increased *RPL23aA* and *B* transcript levels following IAA and BAP treatment while ABA treatment resulted in a transient lowering of transcript levels. Expression patterns differed between *RPL23aA* and *B* in cold-, wound-, and copper-stressed seedlings. In all tissues examined, *RPL23aB* transcript levels were consistently lower than those of *RPL23aA*. This research shows differential transcription of the two *RPL23a* genes, which should no longer be identified as 'housekeeping' genes, and suggests different regulatory mechanisms controlling *RPL23aA* and *B*.

3.1. Introduction

Ribosomes, the ribonucleoprotein particles responsible for peptide synthesis in all living organisms, are among the largest enzymatic complexes of the cell. The eukaryotic ribosome has two subunits (40S and 60S) consisting of four ribosomal RNA

(rRNA) molecules complexed with over 70 ribosomal proteins (r-proteins); all of these constituents are present in equimolar amounts. The number of r-proteins in cytoplasmic ribosomes varies between species; the yeast (*Saccharomyces cerevisiae*) ribosome has 78 identified r-proteins (Mager et al., 1997), while the rat (*Rattus norvegicus*) ribosome is comprised of approximately 80 r-proteins (Wool et al., 1995), as is the human (*Homo sapiens*) ribosome (Uechi et al., 2001). Plant cytoplasmic ribosomes examined to date are estimated to contain between 75 and 92 r-proteins, depending on the species (Bailey-Serres, 1998). A survey of the complete *Arabidopsis thaliana* genome identified 80 cytoplasmic r-proteins (32 of the 40S subunit, 48 of the 60S subunit), based on sequence homology to known rat r-proteins, encoded by 249 genes (Barakat et al., 2001).

Eukaryotic r-protein genes are usually present in multi-gene families. Plants such as rice (*Oryza sativa*; Wu et al., 1995), maize (*Zea mays*; Larkin et al., 1989), and Arabidopsis (Barakat et al., 2001) appear to have families with lower copy numbers than mammals such as the rat; of 59 rat r-protein genes, the average copy number is 12 (Wool et al., 1995). However, unlike the case in mammals, where a large r-protein gene family generally contains one functional member and multiple pseudogenes (Wool et al., 1995; Zhang et al., 2002), plant r-protein gene families usually have multiple expressed members (Wu et al. 1995; Barakat et al., 2001). Arabidopsis r-protein genes are present in multigene families with two to seven members, with an average copy number of three (Barakat et al., 2001). The function of multiple r-protein isoforms remains unclear; it is unknown whether multiple isoforms function in the same role in the ribosome and serve to accommodate increased translational needs, or whether different isoforms play different roles in the ribosome or in the cell. An investigation of the developmental regulation of Arabidopsis r-protein gene *RPL16* (*RPL11* in the nomenclature of Barakat et al., 2001) showed that while the upstream regulatory elements of *RPL16B* drove reporter gene expression in vegetative and floral meristematic tissues, *RPL16A* upstream sequence directed very specific expression behind the root meristem, in lateral root primordia, the stele, and in developing anthers and pollen (Williams and Sussex, 1995). As well as functioning within the ribosome some r-proteins function outside the ribosome, e.g. in DNA replication and repair, transcription, and RNA processing (Wool, 1996). It has been noted that the involvement of r-proteins in extraribosomal activities

supports the idea that r-proteins were recruited to the ribosome from other functions in the cell over time, adding to a rRNA core (Wool et al. 1995; Wool, 1996). That the enzymatic (peptidyl transferase) activity of the ribosome is a function of large subunit rRNA (Nissen et al., 2000; Jenni and Ban, 2003) supports the rRNA-first view of ribosome evolution.

While r-protein genes in *Escherichia coli* are arranged in operons (Mager, 1988) most eukaryotic r-protein genes are dispersed throughout the genome (Mager, 1988; Barakat et al., 2001; Uechi et al., 2001), complicating coordinated expression. Much of the research examining r-protein gene expression has focused on general expression patterns within tissues during development. Transcript levels for specific r-proteins in Arabidopsis (Van Lijsebettens et al., 1994; Williams and Sussex, 1995), tobacco (*Nicotiana tabacum*; Marty and Meyer, 1992; Gao et al., 1994; Dai et al., 1996), maize (Larkin et al. 1989; Lebrun and Freyssinet, 1991; Joanin et al., 1993; Chevalier et al., 1996), canola (*Brassica napus*; Bonham-Smith et al., 1992), pea (*Pisum sativum*; Strafstrom and Sussex, 1992; Moran, 2000), petunia (*Petunia hybrida*; Lee et al., 1999), and potato (*Solanum tuberosum*; Taylor et al., 1992) have all been shown to be elevated in meristems and other actively developing tissues. Treatment of plant tissues with phytohormones such as auxins, or mechanical wounding of tissues, has also been found to result in an increase in specific r-protein gene transcript levels (Gantt and Key, 1983, 1985; Van Lijsebettens et al., 1994; Gao et al., 1994; Dai et al., 1996). Auxin in particular has been found to have a marked effect on r-protein gene expression. Gantt and Key (1983) observed an 8-fold increase in the level of translatable r-protein mRNAs upon treatment of soybean seedlings with (2,4-dichlorophenoxy) acetic acid (2,4-D).

In order to investigate differences in expression between r-protein isoforms, we conducted a study of transcript abundance for the two genes encoding RPL23a in Arabidopsis. Arabidopsis RPL23a was initially identified by sequence identity with rat L23a sequenced by Suzuki and Wool (1993), and confirmed as a member of the conserved L23/L25 family of primary ribosomal RNA (rRNA) binding proteins via complementation of a yeast *l25* mutant (McIntosh and Bonham-Smith, 2001). The L23/L25 r-proteins are critical ribosomal constituents, conserved across all domains of life, incorporating prokaryotic (L23) and eukaryotic (L23a; L25 in yeast) homologues.

The L23/L25 r-protein family members directly bind to a conserved site on domain III of 23S or 23S-like rRNA (Buisson and Reboud, 1982; Vester and Garrett, 1984; El-Baradi et al., 1984, 1985, 1987; Jeeninga et al., 1996), playing a key role in the formation of the large ribosomal subunit. Yeast L25, a protein for which Arabidopsis RPL23aA can substitute *in vivo* (McIntosh and Bonham-Smith, 2001; Chapter 2, this volume), has also been shown to play a role in processing of 60S subunit pre-rRNA (van Beekvelt et al., 2001). Crystallographic studies have identified the location of archaebacterial (*Haloarcula marismortui*) L23 near the polypeptide exit tunnel of the ribosome, suggesting a possible role for eukaryotic L23 homologues in protein secretion during contact with the translocon of the endoplasmic reticulum (ER; Nissen et al., 2000). Further studies have confirmed that prokaryotic and eukaryotic RPL23a homologues interact with components of co-translational targeting and translocation pathways (Kramer et al., 2002; Pool et al., 2002; Gu et al., 2003).

Reverse transcription PCR (RT-PCR) is a sensitive technique for investigating transcript levels. Here we report an extensive study of the differential transcription of *RPL23aA* and *B* in response to a broad range of developmental and stress stimuli. In addition to providing a comparison between r-protein isoform genes, these results present a contrast with previously documented transcriptional responses of other plant r-protein genes.

3.2. Materials and Methods

3.2.1. Plant material and seedling cultivation

Arabidopsis thaliana (ecotype Columbia) plants were used in all experiments. For seedlings grown on culture plates or germinated on filter paper, seed was sterilized overnight (18-20 hours) using a vapor-phase sterilization method (Clough and Bent, 1998). Plate-grown seedlings were grown on ½ Murashige and Skoog medium (MS; Murashige and Skoog, 1962) containing 15 gL^{-1} sucrose and 6 gL^{-1} Phytagar (Gibco Invitrogen, California) on vertically oriented 100 x 15 mm square plates (BD Falcon, New Jersey). Plants grown in soil, on MS medium, or germinated on damp filter paper, were all grown at 23°/18°C, 16 h/8 h photoperiod, 50 μmol m^{-2} s^{-1} unless otherwise noted. All tissues were snap frozen in liquid nitrogen following collection. At least three biological replicate experiments were conducted for each treatment.

3.2.2. Treatments

3.2.2.1. Wild type (untreated)

Untreated tissue was collected from five-week-old soil-grown Arabidopsis plants. Tissues included root, leaf, stem, bract, bud, flower, elongating carpels, and green siliques (fully elongated, no floral organs attached).

3.2.2.2. Phytohormones

Seven to ten day-old plate-grown seedlings were treated with 10^{-3} M indole-3-acetic acid (IAA), 6-benzylaminopurine (BAP), (±)-cis,trans-abscisic acid (ABA), or gibberellic acid (GA_3; all phytohormones, Sigma, Missouri). Treatment of the seedlings was performed essentially as in Williams and Sussex (1995). Seedlings were submerged in phytohormone solutions or a water control for 15 minutes then rinsed twice with sterile distilled water. Tissue was collected at 0 (immediately following rinsing), 4, and 24 h post-treatment.

3.2.2.3. Temperature stress

Seven- to ten-day old plate-grown seedlings were used for all temperature stress experiments. Growth and recovery were carried out at 23°/18°C, 16 h/8 h day/night cycle. Temperature stresses were carried out in a separate growth chamber, 16 h/8 h day/night cycle. Heat-stressed seedlings were subjected to a 32°C heat stress for 1 h then allowed to recover for up to 24 h. Tissue was collected preceding, during, and following heat stress at 0, 0.25, 0.5, 1, 1.5, 2, 4, and 24 h. Cold-treated seedlings were subjected to either 5°C or 15°C for 24 h then allowed to recover for up to 4 h. Tissue was collected at 0, 0.5, 1, 4, 6, 12, 24, 24.25 (15 min recovery), and 28 h (4 h recovery).

3.2.2.4. *Wounding*

Three-week old soil-grown Arabidopsis plants were subjected to wounding by scoring basal rosette leaves once with a razor blade. Care was taken to score one leaf per plant and to maintain the integrity of the scored leaves, keeping them intact. Three or four leaves, each from a different plant, were collected at each time point. Leaves were sampled at 0, 5, 10, 15, 30, and 60 min post-wounding.

3.2.2.5. Copper sulfate stress

Surface-sterilized seeds, distributed on damp filter paper, were stratified at 4°C for four days then allowed to germinate. After 24 h of germination, 2 mL of a water

control or 10 μM, 50 μM, 100 μM $CuSO_4$ solution was applied to each plate and seedlings were collected up to 10 h post-treatment. Three plates of germinating seed (approximately 50 mg) were collected per time point (0, 1, 3, 5, 8, and 10 h).

3.2.3. RNA isolation and RT-PCR

Total RNA was isolated from 50 -100 mg of frozen tissue per sample using the RNeasy Plant Mini kit (QIAGEN, California) according to the manufacturer's instructions. RNA was stored in RNase-free water and diluted in 10 mM Tris, pH 7.5 for quantification via UV spectrophotometry (GeneQuant II, Pharmacia Biotech).

Relative quantitative RT-PCR was performed using a OneStep RT-PCR kit (QIAGEN, California) according to the manufacturer's instructions with the exceptions of primer concentrations (discussed below) and a reduction (1 μL instead of 2 μL) in the amount of Enzyme Mix used in each reaction. All RNA template stocks (4 ng/μL, to a maximum of 1 μg) were treated with 5 U DNaseI (Amersham Biosciences, New Jersey) for 10 min at 37°C to eliminate DNA contamination prior to RT-PCR. An RNase-treated control (template RNA treated with 10 ng RNaseA (USB, Ohio) for 10 min at 37°C prior to RT-PCR) was included in every set of reactions. RNA template concentration was optimized to produce non-saturated product bands; 64 ng of total RNA was used in all reactions.

All reactions were duplexed with gene specific primers (Table 3.1) for the gene of interest (*RPL23aA*, *RPL23aB*, *COR15A*, or *HSP101)* and a primer/competitive primer (competimer) combination to amplify the 18S internal standard (Sung et al., 2001). The 18S primers and competimers have identical sequence (provided by Fatma Kaplan and Charles Guy, U. Florida, Gainesville) with the competimers terminating with a 3' dideoxynucleotide. Primer to competimer ratio was optimized to a final ratio of 2:8 to give non-saturated product bands. Primers for *RPL23aA* (At2g39460) were designed from a BAC sequence (F12L6, Genbank accession no. AC004218), primers for *RPL23aB* (At3g55280) were designed from a BAC sequence (T26I12, accession no. AL132954), *COR15A* (At2g42540) primers were designed from a mRNA sequence (accession no. X64138), and primers for *HSP101* (At1g74310) were designed from a mRNA sequence (accession no. U13949; Young, 2003). Final concentration for *RPL23aA* or *B*-specific primers, 18S primers, and 18S competimers in each 50 μL

Gene Name	Primer Name	Oligo Sequence (5'-3')	Amplified Fragment Length	Conc. in Reaction
RPL23aA	L23A1F L23A1R	CGTGTGTGAAGAATCATTTCAAGCC GCCTCACGTAAGCCTTCTTGG	437 bp	0.2 μM
RPL23aB	L23A2F L23A2R	GGGTTTCTGTTTCGCCGCTCAG CACAGAGCGACAATGATCAGATTAG	535 bp	0.2 μM
HSP101	HSP101F HSP101R	AATCGAAGATGAATCCAG TTGATCACTCTTTCAGCA	213 bp	0.2 μM
COR15A	COR15F COR15R	GGCGATGTCTTTCTCAGGAGC CGGTGACTGTGGATACCATATC	607 bp	0.2 μM
18S	cg359F cg360R (primers)	GGAGCGATTTGTCTGGTT TGATGACTCGCGCTTACT	309 bp	0.01 μM
18S	cg361F cg362R (competimers)	GGAGCGATTTGTCTGGTT TGATGACTCGCGCTTACT	N/A	0.04 μM

Table 3.1. Oligonucleotide primers used for RT-PCR. F, forward primer; R, reverse primer.

reaction was 0.2 μM, 0.01 μM, and 0.04 μM, respectively. A 30 min reverse transcription step at 50°C was followed by heat-inactivation/HotStarTaq activation at 95°C for 15 min and 25 cycles of PCR at 94°C (1 min for the first cycle, 30 s for subsequent cycles), 52°C (30 s), and 72 °C (30 s). All steps were carried out in a PTC-100 thermal cycler (MJ Research). Sequences of amplified DNA were confirmed via automated sequencing (Plant Biotechnology Institute, National Research Council of Canada, Saskatoon).

RT-PCR-amplified products were visualized on ethiduim-bromide stained gels using the Gel Doc 2000 gel documentation system (Biorad). Gel Doc 2000 Quantity One software was used to calculate average band density measurements, which were recorded and used in graphical analyzes. The ratio of target gene product band density to 18S internal control band density was calculated and graphed using Microsoft Excel. Standard error (SE) was calculated from data collected in three to four independent biological replicates. Statistical analysis of RT-PCR data was carried out using SAS version 8.2 for Windows (SAS Institute, Inc., Cary, NC, USA). Data from wild type untreated, hormone, and copper sulfate experiments were analyzed within separate mixed models where treatment effects (tissue type, hormone, copper sulfate concentration) and time points were considered fixed and experimental replicates were considered random. Data for different genes (*RPL23aA* and *B*, *HSP101*, and *COR15A*) were combined into a single ANOVA for each of the above models to allow for a quantitative assessment of the interactions of genes with all fixed effects. Wild type tissue, temperature and wounding stress experiments were analyzed within repeated measures mixed models using a compound symmetry covariance structure that was determined to be the most appropriate by SAS model fitting criteria (i.e. Akaike's Information Criterion, AIC, and Bayesian Information Criterion, BIC). Orthogonal contrasts (one degree of freedom) were used to compare between levels of fixed effects. The denominator degrees of freedom used to calculate the significance of fixed effects were corrected for small sample size using the Kenward-Roger method (Kenward and Roger, 1997). Differences between fixed effects were considered significant at $p \leq 0.05$.

3.3. Results

3.3.1. RPL23aA *and* B *sequence analysis*

Two genes encode RPL23a in Arabidopsis; *RPL23aA* is found on chromosome II and *RPL23aB* maps to chromosome III. Both genes exhibit the same overall structure, with two introns and three exons all approximately the same length (Figure 3.1). A comparison of *RPL23aA* and *B* open reading frames shows 71% identity at the nucleotide level, resulting in 94% amino acid sequence identity between the predicted proteins. A primary sequence comparison of the regions upstream of the open reading frames of each gene, however, shows little sequence similarity. Only 42% and 43% identity is shared between the 1 kb and 500 bp regions upstream of the translation start codons of each gene, respectively, and only 54% identity is shared between the 100 bp immediately upstream of the *RPL23aA* and *B* open reading frames. A 1 kb section of upstream regulatory region was chosen for initial study because this distance approximately encompasses the region between the ATG start codon of each *RPL23a* gene and the 3' end of their respective preceding genes on chromosomes II and III.

The Plant Cis-acting Regulatory DNA Elements (PLACE) database (http://www.dna.affrc.go.jp/htdocs/PLACE; Higo et al., 1999; Prestridge, 1991) was used to scan for putative regulatory motifs in the *RPL23aA* and *B* upstream regulatory regions. While many of the same types of motifs were identified in both upstream regulatory regions, arrangement of these putative regulatory motifs differs between genes (Figure 3.1). Putative motifs found in the upstream regulatory regions of both *RPL23a* genes include an auxin-responsive element (AuxRE, $^{5'}$TGTCTC$^{3'}$; Guilfoyle et al., 1998), found 130 bp upstream of the *RPL23aA* ATG start codon, and 171 bp upstream of the *RPL23aB* ORF. A root-specific element ($^{5'}$ATATT$^{3'}$; Elmayan and Tepfer, 1995) is found 419 and 838 bp 5' to the *RPL23aA* ATG start codon, and 369 and 939 bp in the corresponding *RPL23aB* 5' regulatory region. *RPL23aB* has a cluster of three putative gibberellic acid (GA)-regulatory motifs not found in *RPL23aA*; these motifs are found between 735 and 909 bp upstream of the *RPL23aB* ATG start codon and include a $^{5'}$CAACTC$^{3'}$ regulatory element (CARE; Sutoh and Yamauchi, 2003), pyrimidine box ($^{5'}$TTTTTTCC$^{3'}$; Cercós et al., 1999), and a GA response element (GARE, $^{5'}$TAACAAA/G$^{3'}$; Skriver et al., 1991).

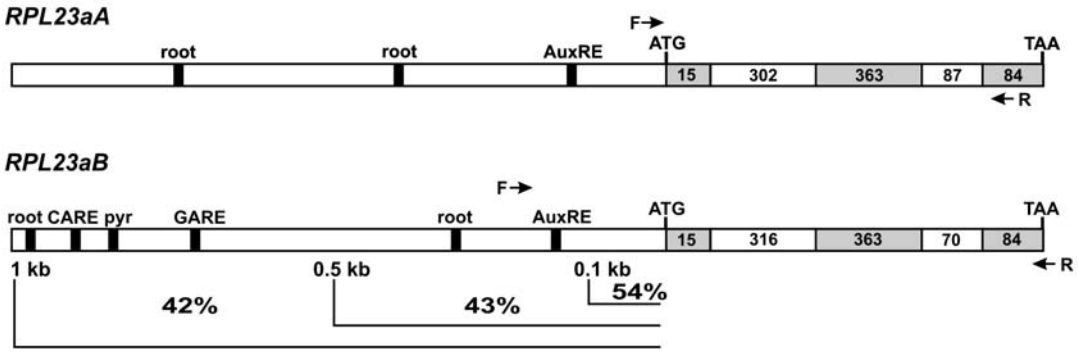


Figure 3.1. Schematic of *RPL23a* genes showing open reading frame and 1 kb of sequence upstream of the ATG start codon. Exons within the open reading frame are shown in grey, introns and 5' upstream region are shown in white. Numbers in exons and introns of the open reading frame indicate length of segment in base pairs. Percentages indicate identity between different regions (0.1, 0.5, and 1 kb upstream of ATG) of *RPL23aA* and *B* 5' regulatory sequence. Black bars indicate putative regulatory elements: root, root-specific element ($^{5'}$ATATT$^{3'}$; Elmayan and Tepfer, 1995); CARE, $^{5'}$CAACTC$^{3'}$ regulatory element (Sutoh and Yamauchi, 2003); pyr, pyrimidine box ($^{5'}$TTTTTTCC$^{3'}$; Cercós et al., 1999); GARE, GA response element ($^{5'}$TAACAAA/G$^{3'}$; Skriver et al., 1991); AuxRE, auxin-responsive element ($^{5'}$TGTCTC$^{3'}$; Guilfoyle et al., 1998). Positions of gene-specific primers used for RT-PCR amplification of each gene indicated by arrows; F, forward primer; R, reverse primer.

3.3.2. Optimization of RT-PCR

Relative quantitative RT-PCR was performed in order to give a relative measure of r-protein transcript compared to an 18S rRNA internal standard. All reactions were duplexed, amplifying the gene of interest (*RPL23aA*, *RPL23aB*, *COR15A*, or *HSP101*) and the 18S internal control. Because rRNA can comprise over 80% of total cellular RNA, amplification of rRNA transcripts would result in a strong, saturated signal. To reduce the strength of the 18S transcript signal, the 18S internal standard was amplified using a combination of primers and competitive primers (competimers, identical in sequence to the 18S primers but with a 3' terminal dideoxynucleotide).

Optimization was carried out as in Sung et al. (2001); template concentration (Figure 3.2) and internal standard (18S rRNA) primer to competimer ratio were both optimized in order to give non-saturated product bands over 25 cycles of PCR following reverse transcription. Four, 16, 32, 64, and 256 ng DNase I-treated total bud RNA were used in optimizing template concentration. A standard template concentration of 64 ng was chosen for all subsequent RT-PCR. Primer to competimer ratios of 2:2, 2:4, 2:6, 2:8, and 2:10 were tested to determine the optimal ratio to produce non-saturated 18S bands (data not shown); a final ratio of 2:8 was selected for all reactions.

RT-PCR was conducted on three to four separate biological replicates for each experiment. Although variation was seen in levels of transcript abundance for each gene examined between replicate experiments, the same overall expression patterns were seen in all replicates (e.g. Figure 3.3). The standard error shown in each graph of relative transcript abundance reflects this variation between replicates.

3.3.3. RPL23aA *and* B *differential expression in untreated plants*

Relative transcript abundance for the *RPL23a* genes was determined in a variety of tissues from plants at stages of development ranging from emerging seedlings, through bolting, to mature plants. Transcript levels were determined for *RPL23aA* and *B* in untreated soil-grown five-week old wild type Arabidopsis Col-0 plants; mature plants were used in order to examine expression in a variety of tissue types. Both genes showed significant differences (*RPL23aA*, $p=0.0003$; *RPL23aB*, $p<0.0001$) in transcript levels between tissues. Levels of *RPL23aA* and *B* transcripts were highest in mitotically active tissues: flower, bud, and elongating carpel tissue showed the highest levels of expression

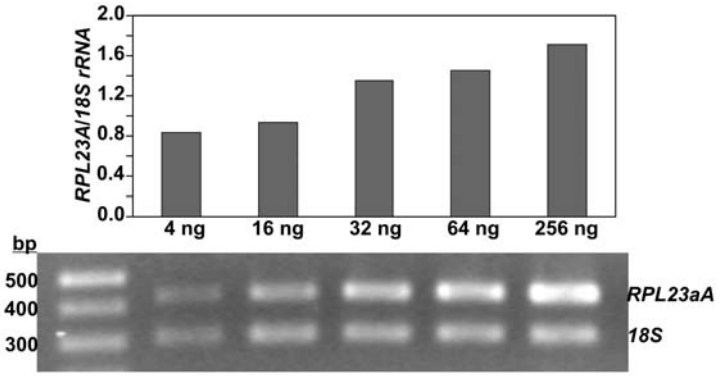


Figure 3.2. RT-PCR optimization for template concentration. Template RNA concentrations of 4-256 ng were used for RT-PCR amplification of *RPL23aA* and *18S* internal control fragments; 64 ng was chosen as the optimal template concentration for all subsequent experiments (non-saturated product band). Band intensities are relative measurements representing the *RPL23aA* to *18S* rRNA average band density ratio.

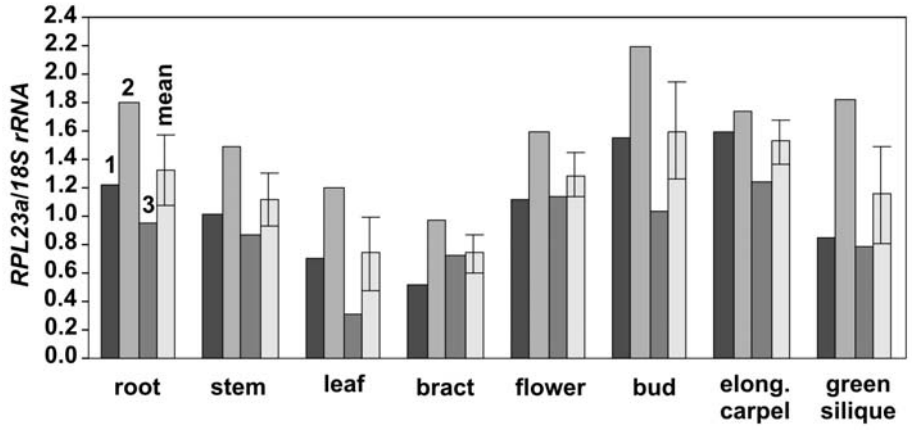


Figure 3.3. RT-PCR amplification of *RPL23aA* using wild type Arabidopsis tissue templates, showing all biological replicates. Band intensities are relative measurements representing the *RPL23aA* to *18S* rRNA average band density ratio. Replicates are labeled 1, 2, and 3; combined replicate average is labeled mean. Means are expressed ± SE (n=3)**.**

for both *RPL23aA* and *B* (Figure 3.4). Root and stem also showed relatively high levels of both gene transcripts (Figure 3.4). While *RPL23aA* and *B* transcripts were present in leaf and bract, they were at significantly lower levels than any of the other untreated tissues examined; *RPL23aA* transcripts were ~1.75 to 2.25 times more abundant in root, stem, bud, flower, and siliques than in leaf and bract (Figure 3.4). *RPL23aB* transcript abundance in leaf and bract was also lower than in other tissues; transcript levels in root and bolt tissues were ~1.5 to 2.4 times greater than in leaf, and ~2 to 3.3 times greater than in bract (Figure 3.4). During maturation of elongating carpel to mature green silique, transcript levels of both *RPL23aA* and *B* were reduced.

3.3.4. Expression in response to phytohormone treatments

Phytohormones are a physiological link between environmental stimuli and genetic response. Given the correlation between *RPL23a* transcript abundance and tissue type/developmental state (Figure 3.4), the effects, on *RPL23a* expression, of some of the phytohormones involved in plant development were examined. In order to gauge the responsiveness of *RPL23aA* and *B* transcription to various phytohormone stimuli, seedlings were treated with 10^{-3} M indole-3-acetic acid (IAA), 6-benzylaminopurine (BAP), (±)-cis,trans-abscisic acid (ABA), or gibberellic acid (GA_3) solutions. Young seedlings were used for all treatments to attempt to maximize sensitivity of the plants to each treatment.

As with untreated tissue, *RPL23aA* and *B* showed similar patterns of transcript abundance in response to phytohormone treatments (Figure 3.5). However, there were quantitative differences in expression between genes, and variation in terms of how transcript levels responded over a 24 h period. Transcript levels for both *RPL23aA* and *B* showed an increase ($p=0.0378$ and $p=0.0126$, respectively) 24 h following IAA treatment (Figure 3.5). BAP treatment also resulted in a significant increase in transcript abundance for *RPL23aA* ($p=0.0114$) and *B* ($p=0.0051$) over a 24 h period (Figure 3.5); *RPL23aA* transcript levels only increased between 4 and 24 h after treatment, while in contrast, *RPL23aB* transcript abundance increased steadily between 0 and 24 h, with approximately one third of that increase occurring in the first 4 h.

Although no ABA-responsive elements were identified in the *RPL23a* upstream regulatory regions by PLACE analyzes, ABA regulates numerous physiological

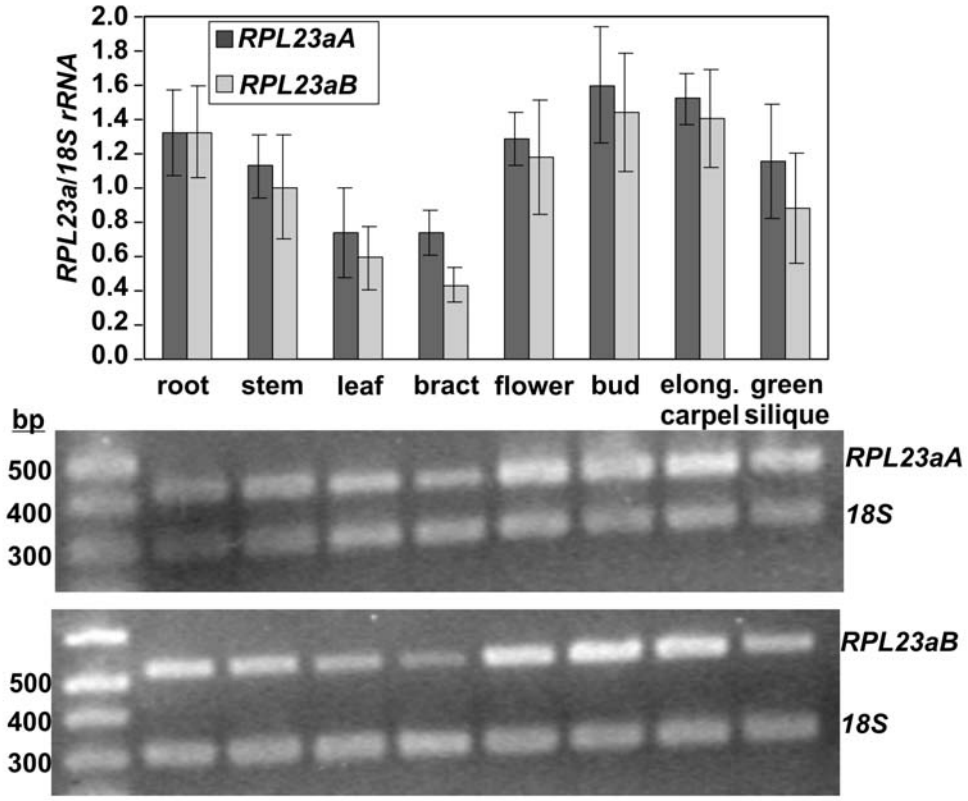


Figure 3.4. RT-PCR amplification of *RPL23aA* and *B* from a variety of wild type Arabidopsis tissues. Band intensities are relative measurements representing the *RPL23aA* or *B* to *18S* rRNA average band density ratio. Means are graphed ± SE (n=3).

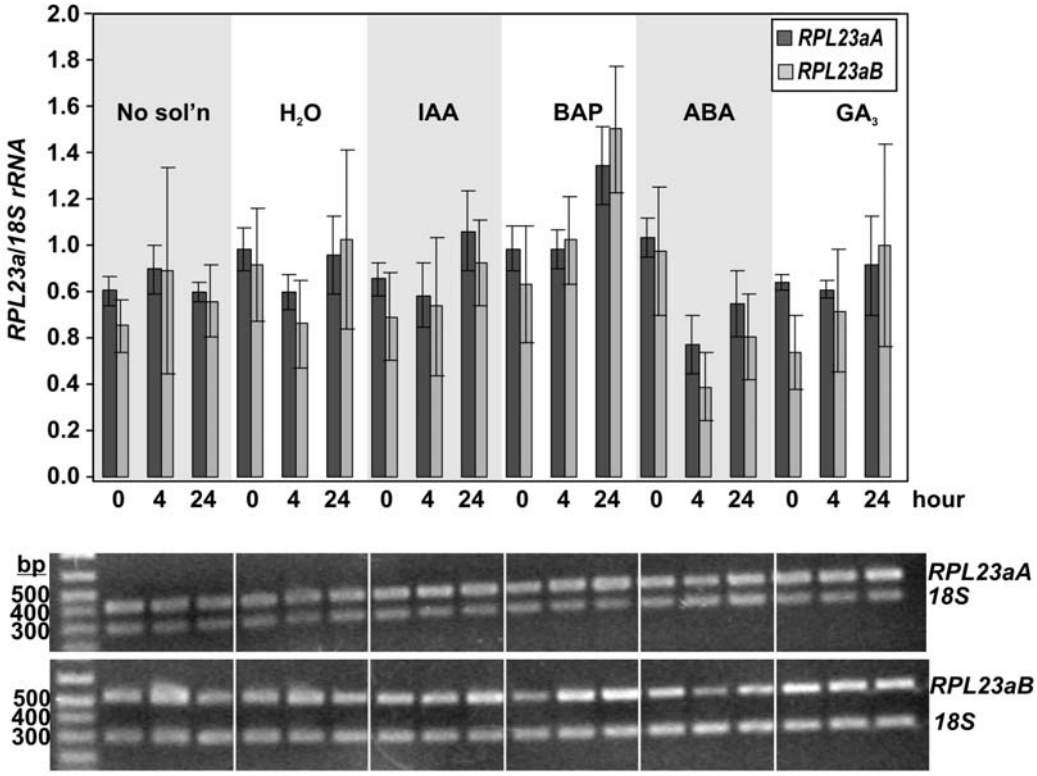


Figure 3.5. *RPL23aA* and *B* transcript levels in response to hormone treatments applied to ~1 week-old seedlings. Seedlings were treated with no solution, water only, or 10^{-3} M IAA, BAP, ABA, or GA_3. Samples were taken at time points 0 (immediately following treatment), 4, and 24 h. White lines in gel photo were overlaid on top of gel image for ease of viewing. Band intensities are relative measurements representing the *RPL23aA* or *B* to *18S* rRNA average band density ratio. Means are graphed ± SE (n=4).

processes throughout plant development (Giraudat et al., 1994) as well as abiotic stress responses (Carrari et al., 2004; Chinnusamy et al., 2004), and as such it was important to determine its effect on *RPL23aA* and *B* transcript levels. ABA treatment resulted in a transient decrease in both *RPL23aA* and *B* transcripts (Figure 3.5), with both genes showing a clear decrease in expression in the first 4 h following treatment (*RPL23aA*, p=0.0024; *RPL23aB*, p=0.0009) and a partial recovery of transcript levels over the next 20 h. *RPL23aB* showed a significantly (p=0.0340) greater decrease in transcript abundance than *RPL23aA* during the first 4 h following treatment, and also showed a greater overall loss of transcript abundance than *RPL23aA* 24 h following treatment.

PLACE analyses of the two *RPL23a* upstream regulatory regions showed the presence of putative GA-responsive elements in *RPL23aB* but not *RPL23aA* (Figure 3.1). The effect of GA_3 on transcript abundance of both *RPL23aA* and *B* was tested to determine if the two genes would show a differential response to the phytohormone. GA_3 treatment affected *RPL23aA* and *B* transcript levels differently; while *RPL23aA* transcript levels remained unchanged following treatment, *RPL23aB* transcript abundance showed an increase over 24 h. Due to the variability between replicate experiments, however, the increase in *RPL23aB* transcript levels was not significant.

3.3.5. Expression in response to temperature stress

Given the critical role of RPL23a homologues in ribosome function and other cellular processes (e.g. van Beekvelt et al., 2001; Kramer et al., 2002; Pool et al., 2002), *RPL23aA* and *B* transcriptional responses to temperature stresses were examined. Known heat- and cold-responsive genes were used as indicators of temperature treatments.

No significant change in transcript levels was detected for *RPL23aA* or *B* after one hour of 32^{o}C treatment (Figure 3.6). Heat treatment was verified by the amplification of *AtHSP101* (*HEAT SHOCK PROTEIN 101*) transcripts; *HSP101* is a known heat-responsive gene (Hong et al., 2001). *HSP101* transcript was produced 0.25 h into the heat treatment and was present up until 0.5-1 h following the heat stress. In contrast, *RPL23aA* and *B* transcript levels remained relatively constant throughout the heat treatment and recovery period. A significant increase (p=0.0363) in *RPL23aB*

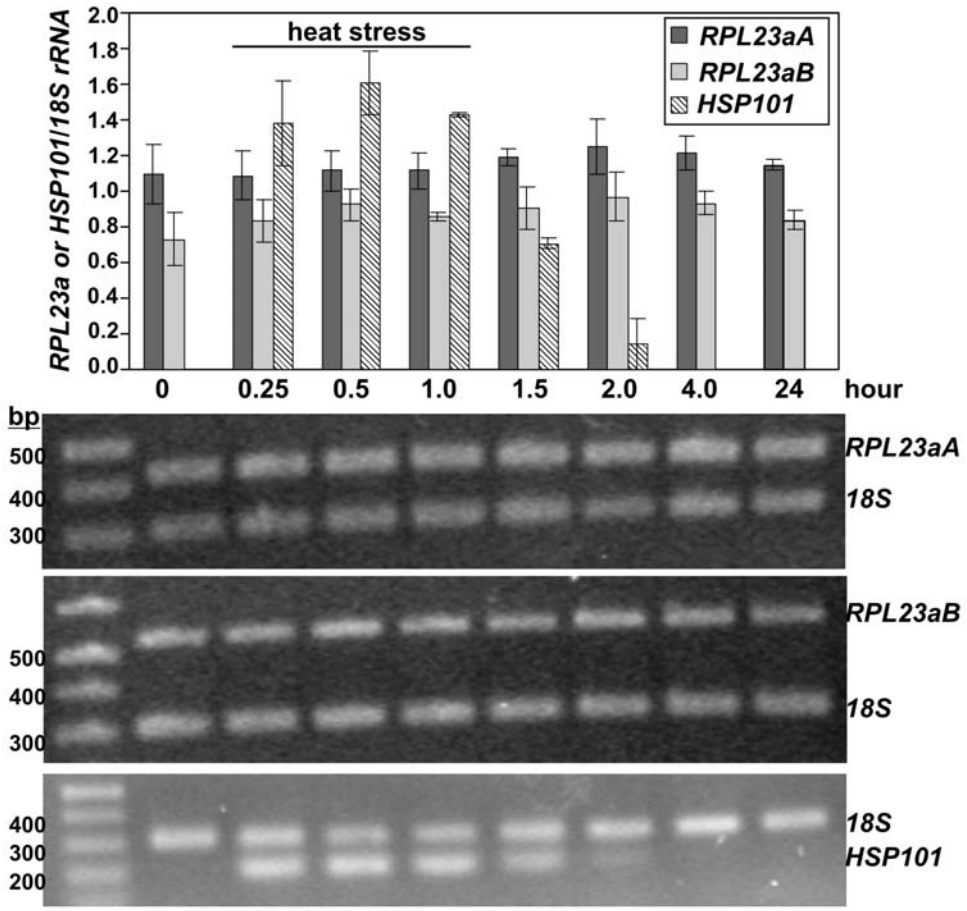


Figure 3.6. *RPL23aA* and *B* transcript levels in response to heat treatment applied to ~1 week-old seedlings. *HSP101* transcript levels were amplified as a positive indicator of heat treatment. Seedlings were treated at 32^{o}C for 1 hour then transferred to normal growth temperature. Band intensities are relative measurements representing the *RPL23aA*, *B*, or *HSP101* to *18S* rRNA average band density ratio. Means are graphed ± SE (n=3).

transcript level was observed one hour following heat stress, but no other significant changes in transcript levels were detected (Figure 3.6).

Cold stress at 5°C for 24 h resulted in no significant change in expression for either *RPL23aA* or *B*, remaining at approximately pre-treatment levels throughout the time course (Figure 3.7). *COR15A*, a *COLD-RESPONSIVE* gene (Lin and Thomashow, 1992), was used as a positive control for the 5°C treatment. *COR15A* transcript was induced between 1 and 4 h into the cold treatment and persisted, at a decreased level, into the first 4 h at the recovery temperature (Figure 3.7). *COR15A* transcript abundance was clearly induced by an effective cold treatment. While *RPL23aB* transcript abundance was more variable between replicates, both *RPL23a* genes showed the same overall stability of transcript levels.

Treatment for 24 h at 15°C had little effect on *RPL23aA* or *B* expression throughout the low temperature time course; however, divergent transcript abundances for *RPL23aA* and *B* were observed during recovery (Figure 3.8). While *RPL23aA* showed an increase in transcript levels over 24 h to the recovery time points, *RPL23aB* showed a sharp decrease in transcript abundance between the end of the 24 h low temperature treatment and the recovery period. Compared to 1 h transcript levels, *RPL23aA* showed an increase by 4 h post-treatment ($p=0.0122$), while the *RPL23aB* transcript level, after 4 h recovery, was lower than at 1 h ($p=0.0068$). Despite nearly identical transcript abundance at the completion of 24 h of 15°C incubation, *RPL23aA* and *B* transcript levels differed by ~50% during the recovery period (Figure 3.8).

3.3.6. Expression in response to wounding stress

Wounding of plants can result from abiotic stress factors such as wind, touch, or rain, or biotic factors such as insect feeding. Plants can respond by upregulating a number of genes that integrate the wound response with a pathogen response. Although no Arabidopsis r-protein genes (including *RPL23aA* and *B*) were identified as being up- or downregulated in a recent microarray analysis of wounded tissue, the analysis was carried out on only 8,200 of the ~22, 500 identified Arabidopsis genes (Cheong et al., 2002). We were interested in determining if either *RPL23aA* or *B* is responsive to a wounding stress. While the overall transcript levels from both *RPL23a* genes in 3-week old untreated and wounded leaves were low (Figure 3.9), corresponding to the levels

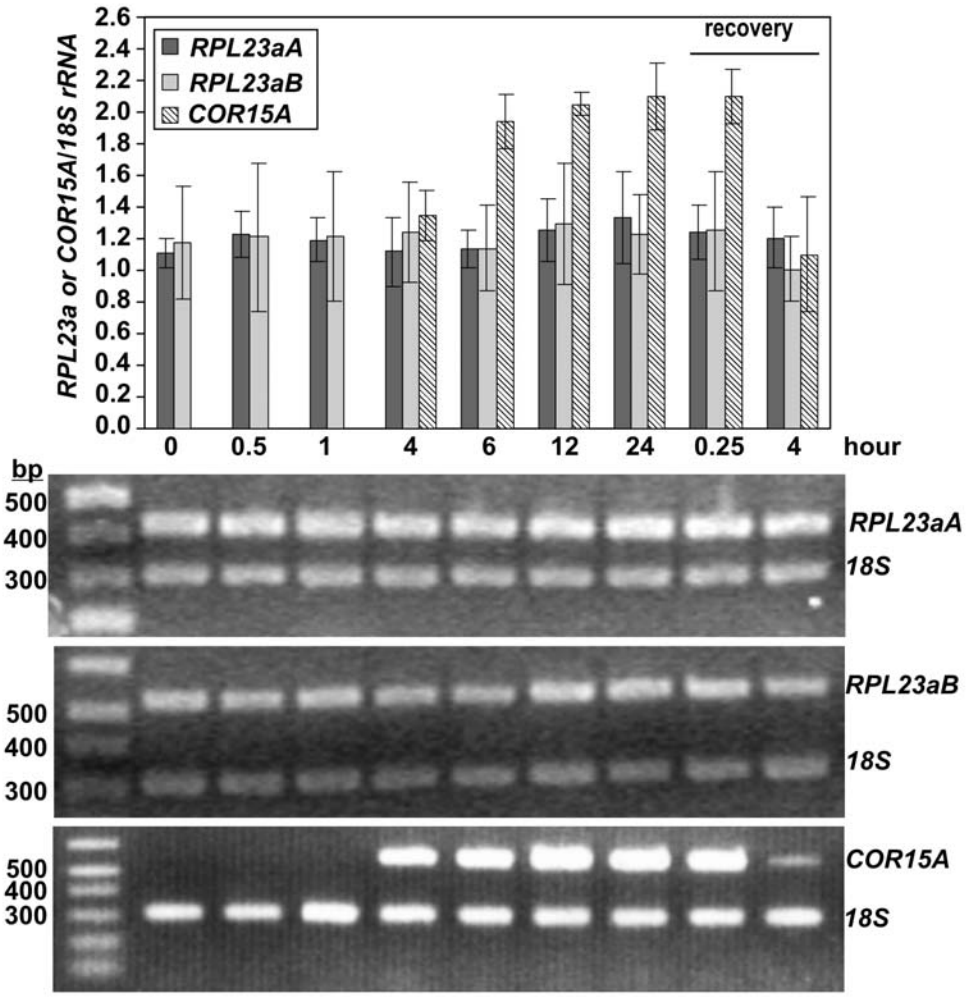


Figure 3.7. *RPL23aA* and *B* transcript levels in response to cold (5°C) treatment applied to ~1 week-old seedlings. *COR15A* transcript levels were amplified as a positive indicator of cold treatment. Seedlings were incubated at 5°C for 24 h then allowed to recover at normal growth temperature. Band intensities are relative measurements representing the *RPL23aA*, *B*, or *COR15A* to *18S* rRNA average band density ratio. Means are graphed ± SE (n=3).

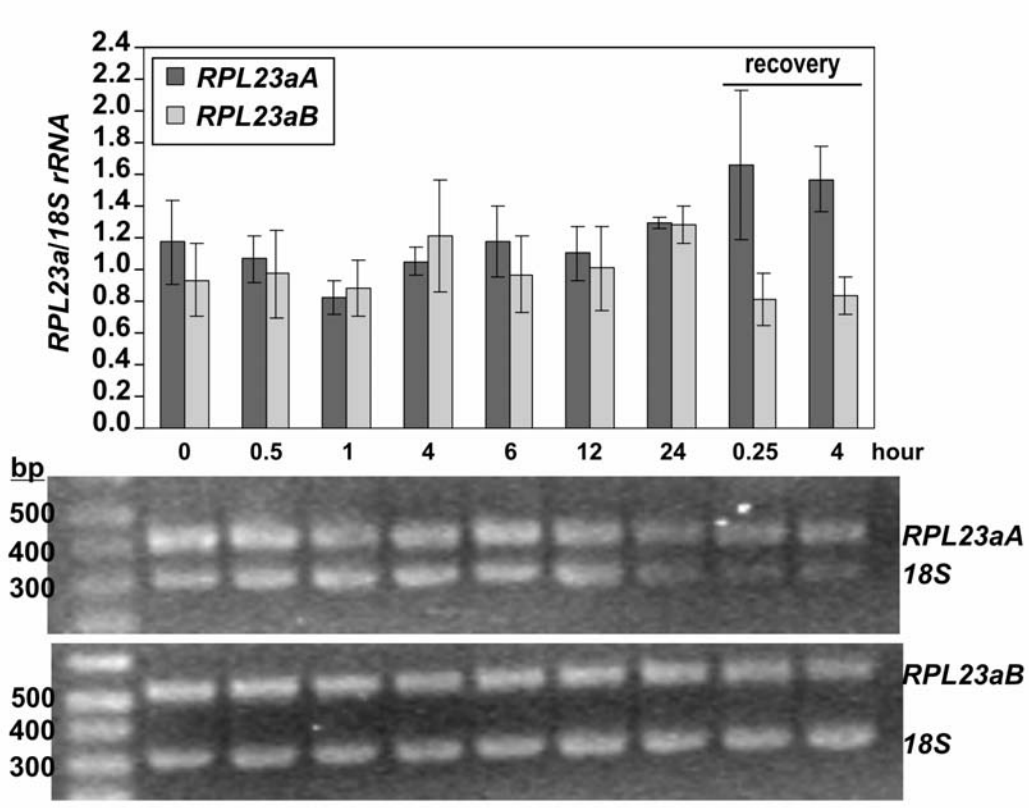


Figure 3.8. *RPL23aA* and *B* transcript levels in response to cold (15°C) treatment applied to ~1 week-old seedlings. Seedlings were incubated at 15°C for 24 h then allowed to recover at normal growth temperature. Band intensities are relative measurements representing the *RPL23aA* or *B* to *18S* rRNA average band density ratio. Means are graphed ± SE (n=4).

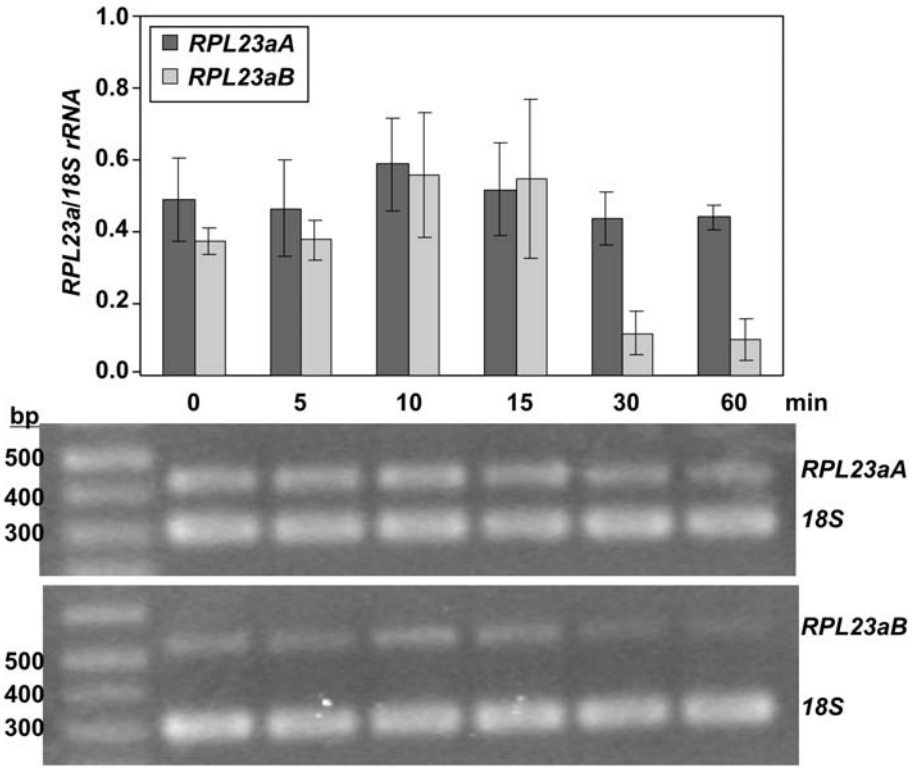


Figure 3.9. *RPL23aA* and *B* transcript levels in response to wounding applied to leaves of 3-week old plants. Leaves were scored on the plants and collected over 1 h following treatment. Band intensities are relative measurements representing the *RPL23aA* or *B* to *18S* rRNA average band density ratio. Means are graphed ± SE (n=3).

previously seen in 5-week old wild type leaves (Figure 3.4), they did show differences in response to the wounding treatment. While *RPL23aA* transcript levels showed no significant differences over the 1 h time course, *RPL23aB* transcript levels increased over 10-15 min post-wounding, then decreased considerably 30-60 min post-wounding (Figure 3.9). While the increase in *RPL23aB* transcript level 10-15 min following wounding was somewhat variable across replicates, the drop in transcript abundance from the 10 to 15 min levels to the 30-60 min levels was consistent across all experiments (10 versus 30 and 60 min, p=0.0233 and 0.0194; 15 versus 30 and 60 min, p=0.0261 and 0.0219). The changes in *RPL23aB* transcript levels following wounding resulted in a significant difference between the two *RPL23a* genes by 30 (p=0.0283) and 60 min (p=0.0063) post-treatment.

3.3.7. Expression in response to copper stress

While wounding and subsequent pathogen attack is a severe stress for plants, many other stress-inducing compounds, including heavy metals, have been shown to elicit a similar response (Ludwig and Tenhaken, 2001). Copper sulfate ($CuSO_4$) treatment was used as a heavy metal stress to see if heavy metal-induced changes in *RPL23a* transcript abundance mimicked that of a wound response. In order to maximize the impact of the $CuSO_4$ treatments, germinating seedlings with radicles just emerging after 24 h of imbibition were used in these experiments. During the 10 h time course, germination continued for all observed seedlings and cotyledons emerged, regardless of treatment, although germination of $CuSO_4$–treated seedlings was slightly retarded. The two *RPL23a* genes showed differential responses to the heavy metal stress (Figure 3.10). While both genes demonstrated a relatively high degree of variability in transcript levels between replicate experiments, $CuSO_4$ treatments had a more significant effect on *RPL23aA* transcript levels than those of *RPL23aB*. While 10 μM $CuSO_4$ treatment had a minimal effect on *RPL23aA* transcript abundance, increasing concentrations of $CuSO_4$ to 100 μM caused a significant (p=0.0311) decline in transcript levels by 5 h which continued to the end of the time course (Figure 3.10). In contrast, *RPL23aB* showed no significant pattern of change in transcript levels during $CuSO_4$ treatments. While 10 μM and 50 μM $CuSO_4$ treatments resulted in a similar pattern of decreased transcript levels as seen for *RPL23aA* following these treatments, *RPL23aB* transcript levels did not

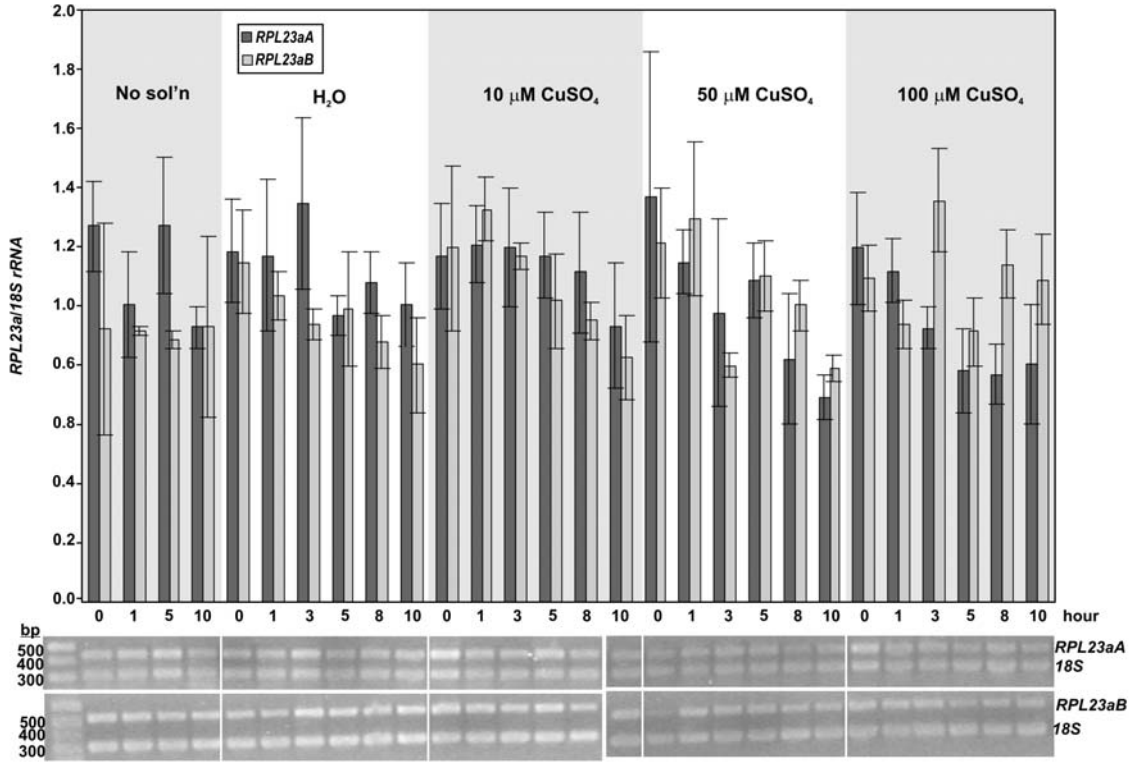


Figure 3.10. *RPL23aA* and *B* transcript levels in response to metal stress. $CuSO_4$ treatments were applied to germinating seedlings after 24 h of imbibition. White lines on gel photos were overlaid on top of pictures for ease of viewing. *RPL23aA* and *B* gel pictures are each composites from two rows run on the same gel due to space constraints. Band intensities are relative measurements representing the *RPL23aA* or *B* to *18S* rRNA average band density ratio. Means are graphed ± SE (n=3).

decline in a straightforward manner (Figure 3.10). Instead of declining steadily over 10 h, *RPL23aB* transcript levels showed a decline between 1 and 10 h (p=0.0226) during the 10 μM treatment and decreased twice during the 50 μM treatment, between 1 and 3 h (p=0.0327) and 1 and 10 h (p=0.0300). No similar pattern of decline in *RPL23aB* transcripts was seen during 100 μM $CuSO_4$ treatment; transcript levels actually increased at 3 h (p=0.0264) relative to 1 h levels during the 100 μM treatment.

3.3.8. Additional differences between RPL23aA *and* B *expression*

In addition to the above noted differences between *RPL23aA* and *B* expression patterns, there were also relative quantitative differences with respect to transcript levels for both genes. *RPL23aA* transcript levels were consistently higher, with only a few exceptions (root, Figure 3.4; 5°C treatment, Figure 3.7; 15°C treatment 24 h time point, Figure 3.8; wounding 10 and 15 min time points, Figure 3.9), than those of *RPL23aB* in all tissues examined under all treatments (Figures 3.4-3.10). *RPL23aB* transcript levels also showed more variance between replicates than *RPL23aA*, as reflected in a higher standard error for most time points (Figures 3.4-3.10); SE for *RPL23aB* transcript levels were an average of 1.5 times higher than those of *RPL23aA* across all experiments. In particular, SE for *RPL23aB* transcript levels for wounding (Figure 3.9), 5°C (Figure 3.7), and phytohormone (Figure 3.5) experiments was an average of ~1.7 to ~2.4 times greater than those of *RPL23aA*.

3.4. Discussion

Ribosomal protein genes are often viewed as a homogenous collection of 'housekeeping' genes with little attention paid to the diversity of this large group. There is a lack of experimental work to support the assumptions that all r-proteins have similar functions or that r-protein genes are all regulated in a similar manner. Not only have many r-proteins been shown to have extraribosomal functions (Wool, 1996), but r-proteins have been routinely identified individually or as part of small groups in screens for genes specifically induced or repressed during different stages of development, in particular tissues, or during stresses; e.g. tuber development (Taylor et al., 1992), root development (Williams and Sussex, 1995), glucose starvation (Chevalier et al., 1996), genotoxic stress (Revenkova et al., 1999), cold stress (Sáez-Vásquez et al., 2000; Kim et al., 2004), heat stress (De Angelis et al., 2004), and depletion of ER

calcium stores (Zhang and Berger, 2004). Despite a high degree of conservation across species and kingdoms, it is becoming clear that r-proteins represent a functionally diverse group with equally diverse means of regulation. Our data suggests that not only are there differences in transcriptional regulation between *RPL23aA* and *B* in Arabidopsis, there are differences in regulation between these *RPL23a* genes and other, previously profiled, r-protein genes.

RPL23aA and *B* transcripts were found in all tissues, from the earliest stages of development examined, after 24 h of imbibition (Figure 3.10), to ~1 week-old seedlings (Figures 3.5-3.8) and mature tissues from 3- (Figure 3.9) or 5-week old (Figure 3.4) plants. A developmental series of seedlings collected from 44 to 122 h during germination also demonstrated this trend, showing the presence of *RPL23aA* and *B* transcripts throughout plant development (data not shown). The presence of *RPL23a* transcripts in germinating seedlings agrees with data from maize that showed r-protein mRNAs are among the stored mRNAs translated in germinating embryos, perhaps even more stable than other stored cytoplasmic mRNAs (Beltrán-Peña et al., 1995). The current data do not allow any conclusion as to whether or not the *RPL23a* genes are actively transcribed during early germination.

Although *RPL23aA* and *B* transcripts were found in all tissues examined, transcript levels were not equivalent in all tissues. As with previously published r-protein data, *RPL23aA* and *B* transcripts were most abundant in mitotically active untreated tissues such as bud, and least abundant in mature, fully expanded leaves (Figure 3.4). Interestingly, *RPL23aA* and *B* transcripts were strongly and equally expressed in root (Figure 3.4); this is consistent with Arabidopsis r-protein L16 (L11 according to the nomenclature of Barakat et al., 2001) data showing an upregulation of expression of two *L16* genes (*L16A* and *L16B*) during lateral root induction (Williams and Sussex, 1995). In addition, auxin application also resulted in a slight increase in transcript levels for both *RPL23a* genes over the 24 h following treatment (Figure 3.5); following a similar exogenous application process, *RPL16-A* and *-B* transcript levels were found to increase in root but not shoot, suggesting that *RPL16* transcription is not directly regulated by IAA but is upregulated during an IAA-induced root developmental process (Williams and Sussex, 1995). Instead of isolating transcript from root and shoot

separately, our study used whole week-old seedlings, and roots were not directly incubated in IAA. If *RPL23aA* and *B* transcription is associated with root-related developmental events instead of being directly affected by IAA, this may have caused the relatively small transcript increase in *RPL23aA* and *B* observed over 24 h.

Other treatments that resulted in similar expression patterns for both *RPL23aA* and *B* included cytokinin and abscisic acid treatments (Figure 3.5). As with results reported for the *RPL23a* tobacco homologue, *L25* (Gao et al., 1994), BAP treatment resulted in an increase in transcript level for both genes, although the increase in *RPL23aB* transcript levels over 24 h was greater than that of *RPL23aA*. While there was an 8-fold increase in tobacco *L25* transcript following cytokinin treatment, this result followed incubation of isolated leaf tissue in MS medium supplemented by the phytohormone, and the wounding caused by isolating the leaf tissue itself caused a 3-fold increase in *L25* transcript (Gao et al., 1994). Our study utilized intact seedlings in order to minimize the impact of factors such as wounding, resulting in less dramatic transcript level increases. In contrast to the results with *RPL23a* and tobacco *L25* genes, *RPS14* transcript levels, examined in lupine (*Lupinus luteus*) and Arabidopsis, were found to increase to a maximum at 5 h following cytokinin treatment, then declined for the rest of a 24 h time course (Cherepneva et al., 2003).

ABA treatment resulted in a decrease in *RPL23aA* and *B* transcript levels by 4 h post-treatment, followed by a recovery to near-0 h levels after 24 h. ABA has also been found to repress transcript levels of other r-protein genes; *RPS14*, *RPS16*, *RPL30*, and *RPL13A* transcript levels in lupine cotyledons all decreased following ABA treatment then increased 24 h post-treatment (Cherepneva et al., 2003). The *RPL13A* and *RPL30* transcripts reached their lowest levels 5 h after ABA treatment then recovered to 60%-90% of their initial levels by 24 h, like the pattern seen with *RPL23aA* and *B*; *RPS16* and *RPS14* transcript levels, however, were much slower to decrease, reaching their lowest levels by 10 – 24 h and not recovering within the 24 h time course (Cherepneva et al., 2003). It is interesting that r-protein genes for large subunit proteins from different species would appear to be regulated in a similar manner in response to ABA while r-protein genes for small ribosomal subunit proteins show a different expression profile.

The effects of GA_3 on r-protein gene expression have not been previously investigated. GA_3 treatment resulted in a variable increase in *RPL23aB* transcript levels between replicates; no such trend was present for *RPL23aA* (Figure 3.5). Interestingly, the *RPL23aB* upstream regulatory region contains a group of GA-responsive elements while *RPL23aA* does not (Figure 3.1), although this did not result in a significant change in *RPL23aB* transcript levels.

Temperature treatments had little effect on *RPL23aA* and *B* transcript levels. Although Volkov et al. (2003) reported a decrease in *RPL23aA* and tobacco *L25* transcript levels during heat stress, no such change was observed in our experiments (Figure 3.6). Indeed, *RPL23aA* and *B* transcript levels across the three replicates of the heat stress experiment showed the least variability of all experiments conducted for this study (Figure 3.6). A possible reason for the discrepancy between our work and that of Volkov et al. (2003) is due to the fact that cut leaves were used for the previously published study, while in this study we used intact seedlings. Temperature treatment at 5^{o}C and 15^{o}C for 24 h, like the heat stress, had little effect on *RPL23aA* and *B* transcript abundance (Figure 3.7; Figure 3.8). Transcript levels for both genes remained at approximately pre-treatment levels throughout the cold stresses much like during the 1 h heat stress treatment. A 5^{o}C treatment of up to 1 week also had no effect on *RPL23aA* or *B* transcript levels (data not shown), suggesting that *RPL23aA* and *B* transcripts are relatively stable during temperature stress. This is in contrast to transcript levels for soybean *GmRPS13*, *GmRPS6*, and *GmRPL37* (Kim et al., 2004) and *Brassica napus RPL13* (*BnRPL13*, initially isolated as *BnC24*; Sáez-Vásquez et al., 2000), which were found to increase following cold (4-5^{o}C) treatment. *BnRPL13* transcript levels increased in etiolated *B. napus* seedlings following 48 h of cold treatment (Sáez-Vásquez et al., 2000), while *GmRPS13*, *GmRPS6*, and *GmRPL37* transcript levels have been found to increase in 3-4-day old soybean plants only after 3-5 days of incubation at 5^{o}C (Kim et al., 2004). While there was no change in *RPL23aA* and *B* transcript levels throughout most of the 24 h 15^{o}C treatment, there was a differential response during the recovery period following treatment (Figure 3.8), with *RPL23aA* transcript levels increasing in the 0.25 – 4 h following 24 h of 15^{o}C incubation and *RPL23aB* transcript levels decreasing during recovery (Figure 3.8). This differential regulation between *RPL23a* genes during

15°C treatment, and between *RPL23a* genes and other r-protein genes during 5°C treatment underscores the complexity of r-protein gene regulation. There are clearly differences in the mode of regulation for each r-protein gene.

Wounding treatment also elicited a differential response from the two *RPL23a* genes; neither gene showed a significant increase in transcript abundance after wounding, but *RPL23aB* transcript levels fell dramatically 30-60 min post-wounding (Figure 3.9). The slight increases observed for *RPL23aA* and *B* transcript levels 10-15 min after wounding did not mirror the 3-fold increase of tobacco *L25* transcript over 5 h following wounding (Gao et al., 1994). This is likely due to the difference in experimental technique, however, since *RPL23a* transcript levels were examined in leaves wounded and left on intact plants while the *L25* transcript levels were observed in isolated leaf pieces cut and incubated in solution (Gao et al., 1994). Similar wounding effects resulting in r-protein gene transcript accumulation have been noted in leaves excised and incubated in solution during studies of phytohormone (Cherepneva et al., 2003) and heat stress (Volkov et al., 2003). The wounding effect should be considered when isolating regulatory factors in r-protein gene expression.

Heavy metal (copper sulfate) stress resulted in another interesting difference not only between the *RPL23a* genes but also between the *RPL23a* genes and soybean *RPL2*. The treatments used to determine *RPL23aA* and *B* transcript levels in response to copper sulfate stress were carried out primarily as in Ludwig and Tenhaken (2001) in terms of $CuSO_4$ concentrations used and time course followed, however we used germinating Arabidopsis seedlings instead of soybean suspension cultures. Transcript levels of soybean *RPL2*, encoding a large r-protein important for translation, were found to undergo a transient downregulation 1-5 h post-treatment, returning to pre-treatment levels by 8-10 h (Ludwig and Tenhaken, 2001). In contrast, *RPL23aA* transcript levels were found to decrease over a much longer time course (>10 h) under the same $CuSO_4$ concentrations, probably due to a decreased rate of copper uptake by the Arabidopsis seedlings relative to that of soybean cell cultures. Dose-dependency of the transcript decrease following $CuSO_4$ treatment, however, was similar between *RPL23aA* and soybean *RPL2*; 100 μM treatment resulted in a steeper decrease in transcript level than

10 μM or 50μM $CuSO_4$ (Figure 3.10). Ludwig and Tenhaken (2001) speculate that *RPL2* downregulation is a mechanism for stress adaptation, removing a key r-protein in order to repress translation and allow rapid turnover of cellular proteins. It is possible that other r-proteins, such as *RPL23aA*, whose homologs are involved in translocon/ribosome interaction at the surface of the ER (Kramer et al., 2002; Pool et al., 2002), respond in a similar manner. The delayed response to heavy metal stress did not allow for a direct comparison between the copper sulfate and wounding treatments. In addition, the high degree of variability in *RPL23a* transcript levels between replicates likely reflects the biochemical variation among germinating seedlings. A study of germinating tomato (*Lycopersicon esculentum*) seeds demonstrated that individuals from a genetically homogenous population varied 100- to 1000-fold in terms of endo-β-mannanase activity, and 9-fold in terms of quantities of reducing sugars, indicating that germinating seedlings can vary greatly in terms of biochemical activity (Still et al., 1997). Although germinating seedlings pooled for the copper stress experiment described here were at approximately the same developmental stage (radicle protruding from the seed coat), the high degree of variability between $CuSO_4$ experimental replicates (Figure 3.10) may have resulted from such individual variation.

Our data, combined with the previous reports discussed above, suggest that r-protein gene regulation is not uniform in nature, but instead varies from gene to gene, even among the most highly conserved r-proteins such as RPL23a. These data challenge the notion that all r-protein genes should be considered ‘housekeeping’ in nature.

CHAPTER 4. SAME FAMILY, DIFFERENTIAL CONTROL: RIBOSOMAL PROTEIN GENES *L23AA* AND *L23AB* ARE REGULATED BY DIFFERENT *CIS*-ELEMENTS

As discussed in the previous chapter, the two genes encoding Arabidopsis RPL23a show differential expression at the transcript level. Transcripts for *RPL23aA* and *B* have been identified in all tissues examined, in both seedlings and mature organs, but transcript abundance differs between the genes both quantitatively and in terms of qualitative response to specific stimuli. Given the relatively low (~40-50%) primary sequence identity between *RPL23aA* and *B* 5' regulatory regions (RRs) and lack of obvious functional regulatory elements for the two genes, an experimental dissection of *cis* elements regulating *RPL23aA* and *B* expression was performed. Transcription initiation site mapping and mRNA amplification showed that the 5' RRs for both *RPL23a* genes are complex, with multiple transcription start sites, and that both genes harbor 5' leader introns that influence gene expression. A series of transgenic plants carrying deletion fragments of *RPL23a* 5' RRs driving *GUS* reporter gene expression demonstrated the differential regulation conferred by the RRs in both full-length and dissected forms, and a region of *RPL23aB* upstream flanking sequence required to direct expression in anthers and pollen was identified. RT-PCR using template from 5' RR deletion series transgenic plants confirmed the importance of post-transcriptional and translational regulation for *RPL23aA* and *B* expression. The following section of this thesis illustrates the complexity of individual r-protein gene expression in plants, further demonstrating differential regulation even between members of the same r-protein gene family. The research also shows how multiple levels of regulation beyond transcription are critical for the control of plant r-protein gene expression.

4.1. Introduction

Responsible for protein synthesis in all living organisms, the ribosome, a peptidyl transferase, is the largest enzymatic complex of the cell. The ~3-4.5 MDa eukaryotic cytoplasmic ribosome is comprised of two subunits (40S, 60S) carrying four separate ribosomal RNAs (rRNAs) and approximately 80 different ribosomal proteins

(r-proteins; Lecompte et al., 2002). The Arabidopsis cytoplasmic ribosome has an estimated mass of ~3.2 MDa (Chang et al., 2005) and is composed of four rRNA molecules (18S, 26S, 5.8S, 5S) and 81 r-proteins, including a plant-specific acidic r-protein (Szick et al., 1998; Barakat et al., 2001), and the RACK1 (receptor of activated C-kinase) homologue in the 40S subunit (Chang et al., 2005). Ribosome biogenesis is linked to growth, development, and stress stimuli, and requires equimolar amounts of all components for correct subunit assembly.

Multigene families encoding each r-protein gene appear to be the rule, rather than the exception, in eukaryotes, but the number of gene family members that are actually expressed varies considerably between species. In mammals, large multigene families encode each r-protein but include only one functional copy among numerous pseudogenes (Wool et al., 1995; Harrison et al., 2002; Zhang et al., 2002). Genome duplications in *Saccharomyces cerevisiae* (Planta and Mager, 1998) and *Xenopus laevis* (reviewed in Amaldi et al., 1995) have resulted in families of two functional genes encoding various r-proteins, although duplicate genes are usually not transcribed at the same level (e.g. yeast, Warner et al., 1985; Jiménez et al., 2002). In contrast, plant genomes contain families of multiple expressed genes encoding each r-protein (Wu et al., 1995; Barakat et al., 2001). Arabidopsis r-protein genes, for example, are present in multigene families of 2 to 7 members with an average copy number of 3 (Barakat et al., 2001). Expression of multiple gene family members for each r-protein may indicate a high translational requirement (Van Lijsebettens et al., 1994), spatial or temporal specificity (Williams and Sussex, 1995), extraribosomal roles for some gene family members (e.g. Wool, 1996), or ribosomal heterogeneity (Chang et al., 2005; Giavalisco et al., 2005).

Given the number of r-proteins in each ribosome, coordination of r-protein gene expression is a complex task in all organisms, and the mechanism(s) by which it is accomplished varies, depending upon the species (Mager, 1988). In prokaryotes such as *Escherichia coli*, the coordination of r-protein gene expression is simplified by the arrangement of r-protein genes in operons (Lindahl and Zengel, 1986; Mager, 1988; Coenye and Vandamme, 2005) and the coupling of transcription and translation in the cytoplasm. Prokaryotic r-proteins are most commonly regulated at the translational level

via feedback inhibition (reviewed in Lindahl and Zengel, 1986; Nomura, 1999); as subunit assembly decreases and free r-protein increases, one r-protein of an operon binds to its own polycistronic message and represses translation (e.g. Nomura et al., 1980; Yates et al., 1980; Raibaud et al., 2003). In eukaryotes, r-protein genes are located throughout the genome (Mager, 1988; Planta and Mager, 1998; Barakat et al., 2001; Uechi et al., 2001) and the presence of a nucleus to separate transcription and translation results in both increased complexity and additional opportunities for control (mRNA processing and transport, r-protein transport and modification) of each process.

In the yeast *S. cerevisiae*, r-protein synthesis is primarily controlled via transcription and is closely linked to changes in nutrient availability, carbon source, and temperature stress (reviewed in Warner, 1989; Li et al., 1999; Warner, 1999). Yeast r-protein genes share a common 5' RR architecture, consisting of two Rap1 (repressor-activator protein 1) binding sites (or an Abf1, autonomously replicating sequence binding factor 1, site) followed by a T-rich enhancer region (Rotenberg and Woolford, 1986; Woudt et al., 1986; Schwindinger and Warner, 1987; Mager, 1988; Warner, 1989; Planta et al., 1995). Rap1 and Abf1 displace nucleosomes around their binding sites, allowing other regulatory factors access to r-protein genes (reviewed in Planta et al., 1995; Lascaris et al., 2000). Factors associated with Rap1/Abf-bound yeast r-protein genes include the histone acetylase Esa1 (Reid et al., 2000), the Rpd3-Sin3 histone deacetylase complex (Rohde and Cardenas, 2003), and the transcription factor Flh1 and its coactivator, Ifh1 (Martin et al., 2004; Schawalder et al., 2004; Wade et al., 2004; Rudra et al., 2005) and corepressor, Crf1 (Martin et al., 2004).

Although up- and down-regulation of r-protein expression in animals is primarily a function of translational regulation, with r-protein mRNAs alternating between active translation in polysomes during growth and development and sequestration in mRNPs in resting or mature cells (Geyer et al., 1982; Pierandrei-Amaldi et al., 1982; Aloni et al., 1992; Loreni and Amaldi, 1992), coordinate expression is also probably aided at the transcriptional level, by a common architecture among vertebrate r-protein genes. At least 60% of mammalian r-protein genes have a canonical or non-canonical (A/T-rich) TATA box at the expected -25 position (Perry, 2005), and a survey of transcription start sites, known to be located within polypyrimidine tracts in other vertebrates (Mager,

1988; Meyuhas and Klein, 1990; Amaldi et al., 1995; see above), yielded the consensus initiator sequence $^{5'}(Y)_2C^{+1}TY(T)_2(Y)_3{}^{3'}$ (Y= any pyrimidine, C^{+1}= transcription start site; Perry, 2005). Initiation within this consensus sequence results in a 5' terminal oligopyrimidine tract (5' TOP) motif characteristic of the 5' UTRs of animal r-protein mRNAs (Amaldi et al., 1995; Meyuhas, 2000). Consensus motifs for the transcription factors YY1 (Yin Yang 1) and GABP (GA-Binding Protein) are each found in over 50% of mammalian r-protein gene 5'RRs, and while common in the r-protein genes of other vertebrates, they are not common to other 'housekeeping' genes (Perry, 2005). Like Rap1 and Abf1 in yeast, GABP and YY1 can act as transcriptional activators or repressors (Genuario and Perry, 1996; reviewed in Thomas and Seto, 1999) and YY1 has been shown to interact with both histone acetyltransferases and deacetylases (Thomas and Seto, 1999). Vertebrate r-protein genes commonly possess introns in either the 5' RR or immediately following the ATG start codon, effectively separating the coding region from the upstream RR (Amaldi et al., 1995; Perry, 2005). A number of *cis*-regulatory elements have been identified in the first introns of mammalian *L7* (Meyuhas and Klein, 1990), *L32* (Chung and Perry, 1989; Chung and Perry, 1993), and *S14* (Tasheva and Roufa, 1995). Introns also play a critical role in the post-transcriptional regulation of r-protein genes via tissue-specific alternative splicing (Xu et al., 1994), or mRNA degradation following inefficient splicing (Amaldi et al., 1989; Chung and Perry, 1989; Aloni et al., 1992; Mitrovich and Anderson, 2000).

In plants, r-protein transcript abundance has been determined in a variety of tissue types in a number of different species, but there is little knowledge of how this abundance reflects regulatory activities. While r-protein transcripts are generally found in all tissue types, r-protein transcript levels in both dicots (Bonham-Smith et al., 1992; Marty and Meyer, 1992; Taylor et al., 1992; Gao et al., 1994; Van Lijsebettens et al., 1994; Williams and Sussex, 1995; Dai et al., 1996; Lee et al., 1999; Moran, 2000; Hulm et al., 2005; McIntosh and Bonham-Smith, 2005; Chapter 3, this volume) and monocots (Larkin et al., 1989; Lebrun and Freyssinet, 1991; Joanin et al., 1993; Chevalier et al., 1996; Dresselhaus et al., 1999; Williams et al., 2003) have been found to be highest in meristems and other mitotically active or developing tissues. Transcript levels for numerous r-protein genes have also been found to increase in response to mechanical

wounding, auxin, and cytokinin treatments (Gantt and Key, 1983, 1985; Gao et al., 1994; Van Lijsebettens et al., 1994; Dai et al., 1996; Cherepneva et al., 2003; Hulm et al., 2005; McIntosh and Bonham-Smith, 2005; Chapter 3, this volume) and decrease in response to abscisic acid (Cherepneva et al., 2003; Hulm et al., 2005; McIntosh and Bonham-Smith, 2005; Chapter 3, this volume) and biotic stress (Ludwig and Tenhaken, 2001).

A few regulatory motifs, which are likely to confer expression in mitotically active tissues, have been identified in plant r-protein genes. The *telo* box motif (plant interstitial *telo*mere motif, consensus $^{5'}$AAACCCTA$^{3'}$; Lenvik et al., 1994; Trémousaygue et al., 1999) has been identified in at least 174 of 216 annotated Arabidopsis r-protein gene upstream RRs, located up-, down- or both up- and downstream of the transcription start site (Trémousaygue et al., 2003), and drives gene expression in root primordia in concert with other *cis* elements such as the *tef* box (Manevski et al., 2000). The *tef* box (*translation elongation factor 1* box), initially identified in the Arabidopsis elongation factor *EF-1α A1* gene (Curie et al., 1991), has also been identified in plant r-protein and other genes that are expressed in mitotically active, cycling cells (Regad et al., 1995; Manevski et al., 1999). The *tef* box motif (consensus $^{5'}$ARGGRYAnnnnnGTM$^{3'}$ where R = any purine, Y = any pyrimidine, and M = A or C) is usually associated with a *telo* box and activates transcription of genes expressed at the onset of the cell cycle (Regad et al., 1995). Another *cis* element found in conjunction with the *telo* box is the *PCNA* (*PROLIFERATING CELL NUCLEAR ANTIGEN*) site II motif ($^{5'}$TGGGCC/T$^{3'}$), identified in 153 of the 174 *telo* box-containing Arabidopsis r-protein gene 5'RRs (Trémousaygue et al., 2003). The site II motif is found almost exclusively upstream of the *telo* box in the Arabidopsis r-protein genes, an arrangement that has also been identified in a sample of 60 rice (*Oryza sativa*) r-protein genes (Trémousaygue et al., 2003). The site II motif, enhanced by the *telo* box, directs expression in actively dividing tissues (Trémousaygue et al., 2003).

Although the *telo*, *tef*, and site II motifs can link r-protein regulation to mitotically active tissues, plant r-proteins also display gene-specific regulation. When r-protein genes or transcripts are isolated during screening for responsiveness to developmental or stress stimuli via microarrays, cDNA libraries, or differential display

(e.g. Berberich et al., 2000; Sáez-Vásquez et al., 2000; Casati and Walbot, 2003; Kim et al., 2004; Toorop et al., 2005), they are isolated individually or in small clusters rather than as the entire ~80 r-protein set, suggesting independent regulation. In addition, members of the same gene family often show expression patterns that differ both spatially and temporally (e.g. Williams and Sussex, 1995; Dresselhaus et al., 1999; Hughes and Friedman, 2005). Given that the equimolar amounts of r-proteins in each ribosome are a result of the sum total of all regulatory activities for each r-protein, varying transcript or transcription levels for specific r-protein genes may simply be a reflection of an emphasis on transcriptional or post-transcriptional regulation for those specific genes. However, differential expression may also reflect roles outside the ribosome (e.g. Wool, 1996) for certain r-proteins. For example, unlike the other three members of its gene family, the Arabidopsis *L7A* gene is transcribed but the L7A protein is not incorporated into ribosomes, suggesting a possible extraribosomal function (Chang et al., 2005). Finally, differential expression may also reflect ribosomal heterogeneity, with gene family members incorporated into different ribosomes, even in a single tissue type (Chang et al., 2005; Giavalisco et al., 2005).

Despite the central importance of ribosomes to the economy of plant cells, little is known about how each of the genes encoding r-proteins is regulated. Previously we and other researchers have reported that members of single r-protein gene families in Arabidopsis are differentially expressed (Williams and Sussex, 1995; Hughes and Friedman, 2005; Hulm et al., 2005; McIntosh and Bonham-Smith, 2005; Chapter 3, this volume). Here we report a detailed study of the upstream RRs of the two members of the Arabidopsis *RPL23a* gene family in Arabidopsis in order to identify important *cis*-elements. We demonstrate the presence of an intron upstream of the coding region in both genes similar to that found in vertebrates, and show that different *cis*-elements drive the differential expression of *RPL23aA* and *RPL23aB*. The importance of post-transcriptional control in the regulation of both *RPL23a* genes is also discussed.

4.2. Materials and Methods

4.2.1. Plant material and cultivation

Arabidopsis thaliana (cv. Columbia-0) was used for wild type tissue and generation of transgenic plants. Seed to be grown on media was sterilized overnight (18-

20 hours) using a vapor-phase sterilization method (Clough and Bent, 1998). Seedlings used for 5' RACE experiments and transgenic selection were grown on ½ Murashige and Skoog (MS) minimal media (Murashige and Skoog, 1965; Sigma, St. Louis, MO), supplemented with 1.5% w/v sucrose, and 0.8% Phytagar (Invitrogen, Carlsbad, CA). Five week old plants for bud, leaf, and other wild type tissue collection were grown in soil. All plants were grown at 23°/18°C 16 h/8 h photoperiod, 50 μmol photons m^{-2} s^{-1}. Tissues used for RNA extraction were snap frozen in liquid nitrogen immediately following collection.

4.2.2. RNA isolation, 5'RACE, and RT-PCR

Total RNA was isolated from ten day old seedlings, and buds and leaves from five week old plants (50 - 100 mg snap-frozen tissue per sample), using an RNeasy Plant Mini kit (QIAGEN, Valencia, CA) according to manufacturer's instructions. Poly(A)$^+$ RNA was isolated from total RNA samples using the PolyAT Tract mRNA Isolation System (Promega, Madison, WI) according to manufacturer's instructions; 60-100 μg total RNA was incubated with 50 U DNaseI (Amersham Biosciences, Piscataway, NJ) for 10 min at 37°C prior to poly(A)$^+$ RNA isolation. Both total and poly(A)$^+$ RNA were used for transcription start site mapping, yielding identical results (data not shown).

Transcription start sites of *RPL23aA* and *B* were determined using a 5' RACE (5' rapid amplification of cDNA ends) system. A 5' RACE kit (Invitrogen, Carlsbad, CA) was used to carry out transcription start site characterization following manufacturer's instructions with the modification that final nested PCR was performed using *Pfu* polymerase (Stratagene, La Jolla, CA). The final nested amplification was performed as follows: 30 cycles of PCR at 94°C (2 min for the first cycle, 30 s for subsequent cycles), 52°C (30 s), and 72 °C (45 s) with a final 10 min extension at 72°C. All steps were carried out in a PTC-100 thermal cycler (MJ Research, Miami, FL). Nested amplification products were blunt-ligated into pBluescript KS$^+$ (Stratagene) at *EcoRV* using T4 ligase (Invitrogen) and sequences of amplified DNA were confirmed via automated sequencing (Plant Biotechnology Institute, PBI, National Research Council of Canada, Saskatoon). Primers used for cDNA synthesis and amplification in 5' RACE are listed in Table 4.1.

Following 5' RACE and mapping of multiple transcription start sites and upstream splice sites for each *RPL23a* gene, reverse transcription polymerase chain reaction (RT-PCR) was used to investigate features of *RPL23aA* and *B* transcripts in various wild type Arabidopsis tissues (root, stem, leaf, bract, flower, bud, elongating carpel, and mature green silique from 5 week old plants). Total RNA from each tissue type was isolated as above. RT-PCR was performed using a OneStep RT-PCR kit (QIAGEN) as in McIntosh and Bonham-Smith (2005; see Chapter 3, this volume), except no 18S internal standard was included. All RNA template stocks (maximum 1 μg RNA) were treated with 5 U DNaseI (Amersham Biosciences) for 10 min at 37^{o}C prior to RT-PCR; 64 ng of total RNA was used in all reactions. The nested 5'RACE primers (GSP3) for each *RPL23a* gene were used as reverse primers in RT-PCR. The forward primers used were designed immediately 3' to mapped transcription start sites. A single forward primer was used for *RPL23aA* reactions, L23aARTintF ($^{5'}$CAGCGGCTTCACCTCTCC$^{3'}$), and two different primers were used to amplify *RPL23aB* transcripts, L23aBRTintF1 ($^{5'}$CCAAGCAACTTGGATC$^{3'}$; 3' to the majority of transcript start sites) and L23aBRTintF2 ($^{5'}$GGGTTTCTGTTTCGCCGC$^{3'}$; 3' to the transcription start site furthest from the ORF).

4.2.3. 5' RR deletion constructs

Constructs were prepared carrying the entire 5' RRs of each *RPL23a* gene (defined as the entire region upstream of each *RPL23a* translation start site, immediately following the preceding gene) or one of a series of truncated 5' RR fragments, ligated upstream of the *uidA* (*β-glucuronidase*, *GUS*) reporter gene. All cloning was carried out using T4 ligase (Invitrogen), all PCRs used *Pfu* polymerase (Stratagene), and all restriction endonucleases were obtained from Invitrogen or MBI Fermentas (Hanover, MD).

Fragments of the *RPL23aA* (At2g39460) 5'RR were PCR-amplified from BAC F12L6 (Arabidopsis Biological Resource Centre, ABRC, Ohio; Genbank accession no. AC004218) and fragments of the *RPL23aB* (At3g55280) 5'RR were amplified from BAC T26I12 (ABRC, Ohio; accession no. AL132954) using a primer series listed in Table 4.2. Once amplified, 5'RR deletion fragments were digested (*Eco*RI, *Bam*HI), purified using a QIAquick PCR purification kit (QIAGEN) according to manufacturer's

Gene	Primer Name	Oligo Sequence (5'-3')
RPL23aA and *B*	AAP	GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG
	AUAP	GGCCACGCGTCGACTAGTAC
RPL23aA	AGSP1	CTGATGAGTGTGTTCACTTTC
	AGSP2	CTTAATCTTCTTCTTGTCAGCACG
	AGSP3	GGCTTGACCAGACTTCACAG
RPL23aB	BGSP1	CTAATGAGGGTGTTGACTTTCTTGG
	BGSP2	GATCTTTTTCTTGTCAGCACGG
	BGSP3	CGCAGGCTTTTTAACGATTTGGCC

Table 4.1. Oligonucleotide primers used for 5' RACE. AAP, Invitrogen 5'RACE Abridged Anchor Primer (forward primer). AUAP, Invitrogen 5' RACE Abridged Universal Amplification Primer (forward primer). All GSPs (Gene-Specific Primers) are reverse primers. GSP1, used for first strand cDNA synthesis; GSP2, used to amplify dC-tailed cDNA with AAP; GSP3, used for nested amplification with AUAP.

instructions, and ligated upstream of *GUS* in pCAMBIA1381Z (CAMBIA, Canberra, AUS) at $^{5'}$*Eco*RI-*Bam*HI$^{3'}$. Full 5' RR fragments were blunt-ligated into pBluescript KS^{+} (Stratagene) at *Eco*RV then digested, purified using a QIAEX II Gel Extraction Kit (QIAGEN) according to manufacturer's instructions, and ligated upstream of *GUS* in pCAMBIA1381Z at $^{5'}$*Eco*RI-*Sal*I$^{3'}$ (*RPL23aA*) or $^{5'}$*Sal*I-*Pst*I$^{3'}$ (*RPL23aB*). All constructs were confirmed via manual (Sanger et al., 1977) or automated (PBI, NRC Canada, Saskatoon) sequencing. *Agrobacterium tumefaciens* strain LBA4404, carrying the pAL4404 (*vir* containing) plasmid (Hoekema et al., 1983), was used as the host for all constructs.

4.2.4. Plant infiltration and transgenic selection

Infiltration of Arabidopsis was carried out using a modified floral dip method. Arabidopsis plants used for transgenic production were grown in pots covered in cheesecloth to prevent soil spillage during immersion in infiltration media. Plants were used for infiltration at approximately 5 weeks post-germination, following the production of secondary bolts. Infiltration media was prepared using *A. tumefaciens* suspended to an OD_{600} of ~0.8-1.2 in 5% sucrose and 0.01% Silwet-L77 as per Clough and Bent (1998). Pots of plants were inverted and immersed in infiltration medium in a vacuum chamber and subjected to 70-100 kPa (~25 mm Hg) vacuum for two minutes. Following infiltration, pots of plants were covered in vented clear bags for three to four days, after which the tops were cut off of the bags and, after another three to four days, the bags were removed and plants were allowed to continue flowering and set seed.

T_1 seed was collected from the T_0 (infiltrated) plants following seed set and plant dry-down. Vapor-phase sterilized (Clough and Bent, 1998) T_1 seed was selected on ½ MS medium supplemented with 25 μg/mL hygromycin (Sigma) to select for the presence of T-DNA inserts and 300 μg/mL of the β-lactamase inhibitor Timentin® (ticarcillin disodium/potassium clavulanate; GlaxoSmithKline, Middlesex, England, UK) to reduce microbial (i.e. residual *A. tumefaciens*) growth. Non-resistant plants turned brown and died at the cotyledon stage following germination under hygromycin selection. Surviving T_1 seedlings (transgenics) were removed from plates and planted in soil at approximately the four leaf stage, then allowed to grow to maturity and set seed.

Gene	Primer Name	Construct	Oligo Sequence (5'-3')	Fragment Size (bp)
RPL23aA	L23aAΔR	All (reverse)	GCGGGATCCGGCTTGAAATGATTCTTCAC	N/A
	L23aA5'FR	A5'FR	GGAGAGGAGGAGCAAATTGTTTACC	1503
	L23aAΔ1	AΔ1	GCGGAATTCGGTAGAAGCCAGTTCAGC	887
	L23aAΔ2	AΔ2	GCGGAATTCGACACGTTTGTATGTTTC	555
	L23aAΔ3	AΔ3	GCGGAATTCGCAACCAAAAGAATCAGTG	396
	L23aAΔ4	AΔ4	GCGGAATTCGGCCCATTTATTCAATCC	295
	L23aAΔ5	AΔ5	GCGGAATTCCCTCTCCAGGTTCGTGTC	145
	L23aAΔ6	AΔ6	GCGGAATTCGCTTCGCTTTCTGGGTTTC	66
RPL23aB	L23aBΔR	All (reverse)	GCGGGATCCTGCTCAAGATAGATTCTTTTC	N/A
	L23aB5'FR	B5'FR	CATGAATTTGAGTTAGAGGATGG	1061
	L23aBΔ1	BΔ1	GCGGAATTCCACTTGATTCACTTGTCATC	650
	L23aBΔ2	BΔ2	GCGGAATTCGTCATTTTCCAATCCTTAAG	563
	L23aBΔ3	BΔ3	GCGGAATTCCGATTTGGACTTTGGTTTG	447
	L23aBΔ4	BΔ4	GCGGAATTCCGATCTAGGGTTTACGG	345
	L23aBΔ5	BΔ5	GCGGAATTCGTGTCATTTTTCTCGTGC	148
	L23aBΔ6	BΔ6	GCGGAATTCGCTAATTGCTATTGCTC	77

Table 4.2. Oligonucleotide primers used for amplification of 5' RR fragments. ΔR primers were used as reverse primers to amplify all fragments.

4.2.5. Characterization of 5'RR activity via GUS detection

Histochemical GUS activity assays, as modified from Sieburth and Meyerowitz (1997), were carried out on 16-19 day old wild type and T_2 seedlings and various tissues from 7-10 week old wild type and T_2 flowering plants. Tissues were collected in microcentrifuge tubes and fixed in 90% acetone on ice for 15-30 minutes, followed by rinsing in a solution of 50 mM $NaPO_4$ (where $NaPO_4$ denotes sodium phosphate buffer comprised of a mixture of monobasic, NaH_2PO_4, and dibasic, Na_2HPO_4, solutions), pH 7.2, 0.5 mM $K_3Fe(CN)_6$, and 0.5 mM $K_4Fe(CN)_6$ for a minimum of 5 min. After rinsing, X-gluc staining solution (50 mM $NaPO_4$, pH 7.2, 2 mM X-gluc [5-bromo-4-chloro-3-indolyl β-D-glucuronide; Rose Scientific, AB, Canada], 0.5 mM $K_3Fe[CN]_6$, 0.5 mM $K_4Fe[CN]_6$) was added and vacuum infiltration of tissues was carried out for two minutes (70-100 kPa/~25 mm Hg). Tissues were incubated in X-gluc staining solution for 24 h at 37°C, after which chlorophyll was removed via incubation in an ethanol series (30%, 50%, 70%, 85%, 95%, 100%; 50% step solution included 5% acetic acid and 3.7% formaldehyde solution). GUS activity was scored on the basis of X-gluc staining visualized under a stereomicroscope (Wild M3Z, Wild Heerbrugg).

4.2.6. RT- PCR amplification of transcripts in transgenic deletion plants

In order to confirm the presence or absence of *RPL23a* transcripts, and any post-transcriptional processing, two-step RT-PCR was carried out using template RNA from T_2 transgenic deletion construct plants. Template RNA was isolated from T_2 seedlings from three independent T_1 lines per construct (*RPL23aA* and *B* 5'FR, Δ4, Δ5, Δ6). Total RNA was isolated from 17 day old wild type and T_2 seedlings using an RNeasy Plant Mini kit (QIAGEN) according to manufacturer's instructions. First strand cDNA synthesis was carried out using SuperScript II RT (Invitrogen) modified as follows from the manufacturer's instructions: total RNA template was treated with DNase I (Amersham Biosciences; ~1U per 100 ng template, 10 min at 37°C) prior to reverse transcription, 200 ng total RNA was used as template for each reaction, reverse transcription was carried out at 50°C for 30 min using a *GUS*-specific reverse primer (pC-GUS-R2, $^{5'}$CCTGGCACAGCAATTGCCCGGC$^{3'}$). Amplification was carried out using Pfu polymerase (Stratagene) with 2 μL (10%) of the first strand synthesis reaction volume as template and the pC-GUS-R2 reverse primer and a gene-specific forward primer.

Specific primers for each construct were as follows; *RPL23aA*5'FR and *RPL23aA*Δ4, L23aA-RTintF ($^{5'}$CAGCGGCTTCACCTCTCC$^{3'}$); *RPL23aA*Δ5, L23aAΔ5 deletion primer (Table 4.2); *RPL23aA*Δ6, L23aAΔ6 deletion primer (Table 4.2); *RPL23aB*5'FR and L23aBΔ4, L23aB-RTintF2 ($^{5'}$GGGTTTCTGTTTCGCCGC$^{3'}$); *RPL23aB*Δ5, L23aBΔ5 deletion primer (Table 4.2); *RPL23aB*Δ6, L23aBΔ6 deletion primer (Table 4.2).

4.3. Results

4.3.1. RPL23aA *and* B *transcription start sites and unusual transcript splicing*

In order to characterize the *cis* regulatory regions of *RPL23aA* and *B*, mapping of the transcription start sites for both genes was carried out via 5' RACE using total and poly(A)$^+$ RNA template from Arabidopsis bud, leaf, and whole 10 day old seedlings. Following first strand cDNA synthesis with gene-specific primers, an oligo-dC tail was added to the cDNA 5' ends, the tailed cDNA was amplified, and the amplification products were used as template in a subsequent PCR with gene-specific nested primers located in exons approximately 90 bp 3' to the ATG start codon. The final (second) amplification using primers for the 5' dC tail and nested primers yielded ~200-300 bp amplification products for both *RPL23aA* and *B* (Figure 4.1); amplification products did not differ between the three tissue types used (data not shown). While the amplification of *RPL23aA* cDNA 5' ends yielded a single band, at least two bands resulted from the *RPL23aB* amplification (Figure 4.1). The nested 5' RACE fragments were cloned, and multiple clones were sequenced, in order to determine the transcription start sites for each gene and to differentiate between amplification products for *RPL23aB*.

A comparison of the *RPL23aA* and *B* 5' cDNA ends with their respective genomic sequences showed that as well as an intron within the ORF of each gene, *RPL23aA* and *B* cDNAs both showed splicing of introns upstream of their ORFs (Figures 4.2 and 4.3). *RPL23aA* has a 107 bp intron 29 bp upstream of the ATG start codon, flanked by canonical ($^{5'}$GT-AG$^{3'}$) splice sites. In *RPL23aB* there are two alternate sets of splice sites ($^{5'}$TC-TA$^{3'}$; $^{5'}$GT-TC$^{3'}$), neither canonical, that remove almost identical intron sequences. The intron spliced at $^{5'}$TC-TA$^{3'}$, identified in four of the *RPL23aB* 5' RACE clones, is 214 bp in length and is found 32 bp upstream of the ORF; the second version of the intron is 215 bp in length and 23 bp upstream of the

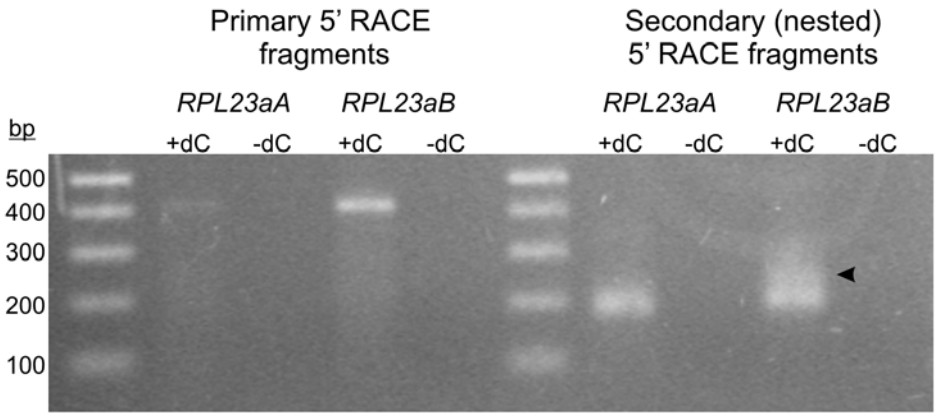


Figure 4.1. 2% agarose gel of *RPL23aA* and *B* 5' RACE products. +dC, amplified product of dC-tailed mRNA; -dC, amplification reaction using non-dC-tailed control template. Arrow indicates second (top) band in *RPL23aB* nested amplification reaction.

A) *RPL23aA*

...ttataga cccaactaag tatcttaggg tttcagtttc cactataaaa cttgttgcgg

cgtaagttag ggttttgaga atcagcggct tcacctctcc aggttcgtgt ctcgattttg

caaactatct ctcgaaatcg tcctacattt ctttcttcag attcattact gagcttcgct

ttctgggttt ctatctaaaa atttcacaga tttcgtgtgt gaagaatcat ttcaagccat

gtctccggct aaaggtacga tcttttcatt gataaccact tgaatcttca ctggaacatg

atggatttgg atcttcttct actagacttt agttttgctg aacccgctag tatctggtct

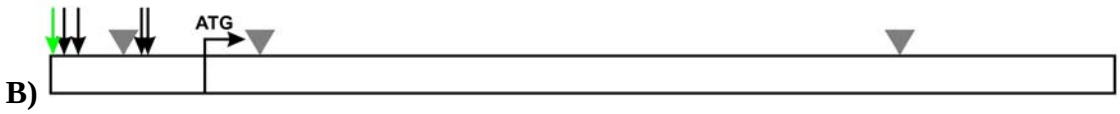
cgattaatgt tttctaatca tgtttcgtta ggtgattagt ttcaatatct tcctgattag

cttcaactaa gtgagattaa tgtagtatcc aaatgaattt tgtctgtaga cgtgtttctg

agtattcatg tattgaagtc agatttgtat caaagtatta tatgatgctt aaactctctg

ttgtatgtaa tttgcagttg atactaccaa gaaggctgat cctaaggcca aggccttgaa

ggcggcaaag gctgtgaagt ctggtcaagc cttcaagaag aaggacaaaa agattag...



B)

Figure 4.2. A) *RPL23aA* transcription initiation and processing sites mapped to genomic sequence, showing portions of coding and 5' RRs. ORF sequence indicated by right-angle arrow at ATG start codon, dots indicate continuation of sequence. Black arrows, transcript start sites determined via 5' RACE; **green** arrow, transcript start site of GenBank cDNA clone; grey shading, introns; underlined sequence, repeat motifs found at upstream intron splice sites; **pink** sequence, gene-specific primer used for 5' RACE nested amplification; **dark blue** sequence, forward primer used for RT-PCR, L23aARTintF. B) Schematic diagram of *RPL23aA* transcripts. ORF is indicated by region following right-angle arrow. **Grey** triangles indicate spliced introns. Black and **green** arrows as in A).

A) *RPL23aB*

...ttgtaat cagttggtag agaaagccca attatatatt atttaaggcc caaaataaat

cgatctaggg tttacggttt tgtttcattt tcacttccct tagaataaat aaaaaccacc

aagcaacttg gatctactct agggtttctg tttcgccgct caggttcgtc aatctctcga

tatcctttct ctaccttctc ctcagctttc tgagtattga atgttgttgt taactgtctc

tgaatacgat tgattttgtg tcatttttct cgtgcttcgt atttgttgat tgattgattt

acatgtcatg atttgaaatc tttgtattgc taattgctat tgctctgttt actgattttc

ttttctgggg ttatcaggtt tcgtgaaaag aatctatctt gagcaatgtc tccagctaaa

ggtacgcctt ttttctttc attgttatct gatttgattc tacatcgtct tctatttctc

ttttgcgtgt aattactttg tttatcatta ttcaatactg tgactgaact tactactgta

tacagctttc tgaaatgtat gattcttatc tggctctgta aagttcgtat tttttatatt

ctcttgctcg tacctactct gttatggatt catctgatta tgtgcttgtt tgtctgttat

ttgccaatga tcaaactttc cttgaatgag atttgttacc agtagtgtaa atggtttaat

gttccattat ttttgcagtt gatgtcacca agaaagccga ccctaaggct aaggctttga

aagctgcgaa agcagtgaaa tctggccaaa tcgttaaaaa gcctgcgaag aagatca...

B)

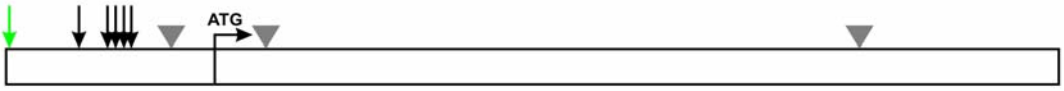


Figure 4.3. A) *RPL23aB* transcription initiation and processing sites mapped to genomic sequence, showing portions of coding and 5' RRs. ORF sequence indicated by right-angle arrow at ATG start codon, dots indicate continuation of sequence. Black arrows, transcript start sites determined via 5' RACE; **green** arrow, 5' transcript start site of GenBank cDNA clone; grey shading, introns; **pink** diamonds, $^{5'}$TC-TA$^{3'}$ splice sites; **lilac** diamonds, $^{5'}$GT-TC$^{3'}$ splice sites; underlined sequence, repeat motifs found at upstream intron splice sites; **pink** sequence, gene-specific primer used for 5' RACE nested amplification; **dark blue** sequence, forward primer used for RT-PCR, L23aBRTintF1; **light blue** sequence, L23aBRTintF2. B) Schematic diagram of *RPL23aB* transcripts. ORF is indicated by region following right-angle arrow. **Grey** triangles indicate spliced introns. Black and **green** arrows as in A).

ATG start codon, and was only identified in one of the 5' RACE clones. A comparison of two *RPL23aB* cDNA clones retrieved from GenBank (accession nos. AY0857361, AY050445) with the *RPL23aB* genomic sequence identified processing at the $^{5'}$TC-TA$^{3'}$ splice sites, identical to the majority of *RPL23aB* 5' RACE clones. A set of direct repeats surrounds the 5' and 3' splice sites of the *RPL23aB* upstream intron(s), $^{5'}$TCAGGTT(T)CGT$^{3'}$, differing only in the presence of an extra T in the repeat surrounding the 3' splice site; the two variations of the *RPL23aB* upstream intron only differ in which repeat is spliced out. Interestingly, similar but more degenerate direct repeats, $^{5'}$CAGGTTCGTGT$^{3'}$ and $^{5'}$CAG(AT)TTCGTGT$^{3'}$, surround the 3' and 5' splice sites, respectively, of the *RPL23aA* upstream intron.

Both *RPL23aA* and *B* show multiple sites of transcription initiation. Transcription start sites as determined by *RPL23aA* 5' RACE fragments extend between 22 and 26 bp 5' to the upstream intron, initiating at A and G nucleotides, respectively. Four *RPL23aA* cDNAs from GenBank which show identical splice sites to the *RPL23aA* 5' RACE fragments, show transcription initiation points at adenine (cDNA accession no. AF325056), guanine (AY037325, AY039850), and thymine (AY086212) 19, 24, and 30 bp 5' to the upstream intron. As a result of differing transcript start sites, *RPL23aA* cDNAs have 5' UTRs between 48 and 59 nt. Two *RPL23aA* 5' RACE products showed transcription initiation sites at pyrimidines (C and T) 25 and 27 bp upstream of the ORF, respectively, and did not include the upstream intron.

RPL23aB transcript start sites as determined by 5' RACE lie between 18 and 42 bp 5' to the most common (214 bp) upstream intron, with most (four of five) transcription initiation sites falling within the first 24 bp upstream of the intron; the resultant 5' UTRs are between 50 and 74 nt in length. One of the *RPL23aB* cDNAs from GenBank (AY0857361) has a 118 nt 5'UTR with a transcript start site 86 bp 5' to the upstream intron; the other GenBank clone (AY050445) has a 51 nt 5'UTR 19 bp upstream of the first intron, identical to one of the 5' RACE clones. Considering both 5' RACE-determined cDNA ends and the cDNA sequences from GenBank, the most common nucleotide found at *RPL23aB* initiation sites is A (four of six cDNAs); two 5' RACE-amplified cDNA ends initiate at a G and a C, respectively.

In order to further characterize *RPL23aA* transcript lengths and intron processing in *RPL23aA* and *B* transcripts, RT-PCR was carried out using RNA templates from a variety of wild type tissues from 5 week old Arabidopsis. Nested 5' RACE primers were used as reverse primers, and forward primers were designed to amplify regions of transcript just downstream of transcription initiation sites as determined by 5'RACE (Figures 4.2A and 4.3A). The primers, upstream of the 5' intron splice sites for the *RPL23a* genes, were able to distinguish between spliced or unspliced transcripts (containing or lacking the upstream intron). A single *RPL23aA* primer, L23aARTintF, was designed to amplify 49 nt of 5' UTR (in 145 bp transcript fragment), and two separate *RPL23aB* primers, L23aBRTintF1 and L23aBRTintF2, were used to amplify 50 or 73 nt of 5' UTR (in 160 or 183 bp transcript fragments), respectively, of processed transcript. Amplification from all wild type tissue templates (root, stem, leaf, bract, flower, bud, elongating carpel, and mature green silique) with the L23aARTintF and L23aBRTintF2 primers produced single bands of expected sizes for the processed *RPL23aA* and *B* transcripts. Sequencing of the L23aARTintF- and L23aBRTintF2-amplified fragments positively identified the transcripts as processed *RPL23aA* and *B* cDNAs lacking upstream introns and spliced at the same sites ($^{5'}$GT-AG$^{3'}$; $^{5'}$TC-TA$^{3'}$) as the previously identified 5' RACE fragments. Amplification with L23aBRTintF1, however, produced no bands using templates from all wild type tissues. The lack of amplification of *RPL23aB* transcripts with longer 5' UTRs indicates that this species of message, isolated once via 5' RACE and identified in a single GenBank clone (AY0857361), is far less common than those mRNAs with shorter (~50 nt) 5' UTRs.

4.3.2. GUS activity driven by serial deletions of RPL23a *5' RRs*

4.3.2.1. Serial deletion constructs and transgenic plants

In order to define the region upstream of each *RPL23a* gene required for gene expression and to identify *cis*-elements important for regulation, serial 5' deletions of each gene were used to drive reporter gene expression in Arabidopsis. The 5' deletion fragments of the upstream flanking regions of *RPL23aA* and *B* were cloned upstream of the *GUS* reporter gene in a binary vector used to transform Arabidopsis, and the resultant plants were screened for the presence of the T-DNA. Following selection of T_1

plants and T_2 seed set, T_2 plants, selected again on media to eliminate homozygous recessive plants (lacking T-DNA), were used for all subsequent assays.

Figure 4.4 depicts the serial 5' deletions of the upstream RR amplified from the BAC for each *RPL23a* gene. Full-length 5' flanking regions were defined as the entire region between the *RPL23aA* or *B* ATG start codon and the annotated 3' end of the preceding gene; for *RPL23aA* this region is 1503 bp, and the full *RPL23aB* 5' flanking region is 1061 bp. A total of seven constructs was assembled for each *RPL23a* gene containing either the full 5' flanking region (designated A or B5'FR) or one of six deletion fragments (designated A or BΔ1, Δ2, Δ3, Δ4, Δ5, or Δ6) upstream of *GUS*. Each deletion fragment was designed to eliminate putative regulatory motifs identified using the PLACE database (Plant Cis-acting Regulatory DNA Elements database, http://www.dna.affrc.go.jp/htdocs/PLACE; Prestridge, 1991; Higo et al., 1999).

4.3.2.2. GUS expression in seedlings

Initial assays of GUS activity were conducted using numerous 16-19 day old T_2 seedlings generated from 5-10 independent T_1 lines for each construct (except *RPL23aB*Δ1, one T_1 line only due to poor transformation efficiency with this construct). Seedlings were scored primarily for the presence or absence of staining for GUS activity (Figure 4.5), and the majority of seedlings that tested positive for GUS activity were mostly or completely blue (++ in Figure 4.5). All wild type seedlings, included as a control, scored negative for GUS activity. While staining was seen in *L23aA*5'FR, AΔ1, AΔ2, AΔ3, and AΔ4 seedlings, AΔ5 (fragment 145 bp upstream of ORF) seedlings showed little to no staining, and staining was completely abolished in AΔ6 seedlings (66 bp fragment 5' to coding region). In contrast, *RPL23aB* transgenics showed at least some staining in all lines, with only BΔ5 seedlings (containing 148 bp upstream of the ORF) showing staining in fewer of the seedlings sampled. BΔ6 seedlings (77 bp 5' to ATG start codon) 'recovered' the staining (all seedlings stained) seen in B5'FR to BΔ4 seedlings.

4.3.2.3. GUS expression in mature plants: different tissue types

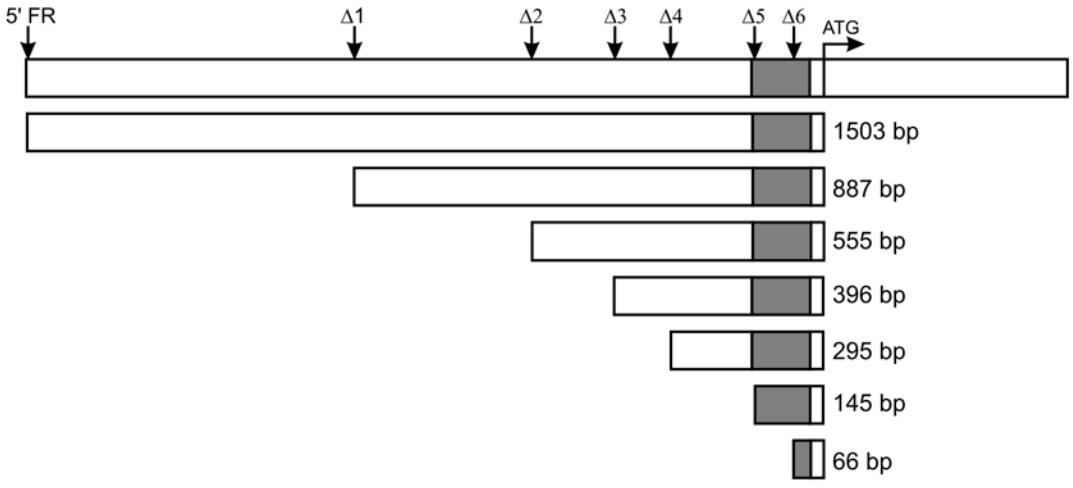
Between one (*RPL23aB*Δ1 only) and five T_1 lines were used to generate T_2 plants for assays of mature plant tissues. Six to nine 16-19 day old T_2 seedlings from

A)

ggaga **5'FR**

ggaggagcaa attgtttacc caaacaacca agacagttgt gtagtttact cggagatttc
ttctccttgt gttccggtga aacctggaaa tcatcgaaat ttaactgcca ggaacggtca
ccgttaccgt gaaactcgag cagcgagatc ccttcgtcgg aatctggact agctggctca
accatttttc cggcggcgga gagagacttt gtatttttt ttttttttt ttgagagaga
gagagagtgg taaggaaatg ggaaatttga agagacggcg agggtaaaat cggtaattaa
gaaagattga ggctagctac tacgactgtt gtggtagtaa gacattcctc cacagagata
cttggaccag cctttagaag agacgtctct gctaattatc cacgtggcgt tctctcaatg
gatttgggat ctcgcacgtg attggaggag tcatcataca cacgtgtaca gttgagattc
aactgcaatt aaacgacccc gtgacatgac atggttatga gtcacatatg tgccctctca
aactctatta tgtagtataa aaactgtatt agatacaacc gacccaagat atagaagaac
ctttgaaaaa tggtagaagc cagttcagct aaagctttct tcttcacaat tatcgtaaaa **Δ1**
tattagacat gtagagcagg cttgttgttc tcttattgtc aagcgtaagt aaatgattat
tgtttagttt cgtgaccaat tggtttgttt tttggttagc cattgttcga tttgctttgg
tttggtcaaa gattgtatgc tttcaagtca aaccaaagct gactaaattt tattctgtaa
taatttactt tggaatggca aaaccaaaga aaggtcatag acacaagaga agaagcaaat
atgtttctta cgccaaaaga cacgtttgta tggtaagaca aaagacacgt ttgtatgttt **Δ2**
cttacgccaa taccattacg ttaaaaataa actagtaact taaatttacg ctcactttct
atttgggttt tacatgatga ttttggtcaa aggcctattt acaatgttgc aagcccatat
tatcttttaa gtatcttttt gtgcaaccaa aagaatcagt gatcttatgc tcattttcat **Δ3**
ttttttctaa tgaaattttt cgtcacttat ggcttatgcg tccgatggac gattatggta
aaaggcccat ttattcaatc caaagcttac tgggttagat gatgattgtg gtaaaaaggc **Δ4**
ccattataga cccaactaag tatcttaggg tttcagtttc cactataaaa cttgttgcgg
cgtaagttag ggttttgaga atcagcggct tcacctctcc aggttcgtgt ctcgattttg **Δ5**
caaactatct ctcgaaatcg tcctacattt ctttcttcag attcattact gagcttcgct **Δ6**
ttctgggttt ctatctaaaa atttcacaga tttcgtgtgt gaagaatcat ttcaagcc...**ΔR**

B)



C)

catgaa tttgagttag aggatggttg gaacaaaaaa acttagaagc tcgaatgacc **5'FR**
ggtttttacc aaattctcat agaccatatt tgattctttt gatttacttc tggtgcagga
ctctctgtgc ttatggaagt tgatgttggg ggaaacaact ctcttgtaca gtggggaaaa
aacttcttct tcttctttct atcacatgaa aatcctcaag ggccattatt agtatgatca
gattataaaa ttgtaaggtt aggggcttta tgaggatttt gatggacttg ttacaatgtt
tacatataca ctcagcagca caatagattt ttgttaaact tacatgttat tcaagtaaaa
gtactatgta gatgttgaag tctaattgaa gaattagtta atgatagtct taaacacttg **Δ1**
attcacttgt catccaattt tggttttgcg catagtttct cttcttttat ttcctctcta
aaacaccaaa accaaacaaa atgtcatttt ccaatcctta aggtttcatt cattttagtg **Δ2**
atttttggg tacaaaattg agcaatgtct agtgacgttt ttactcaaac tcataaacca
acattctaat cagaatcacg atttggactt tggtttggga ccttctttct acaccaactg **Δ3**
ggcttgtaat cagttggtag agaaagccca attatatatt atttaaggcc caaaataaat
cgatctaggg tttacggttt tgtttcattt tcacttccct tagaataaat aaaaaccacc **Δ4**
aagcaacttg gatctactct agggtttctg tttcgccgct **caggttcgtc aatctctcga**
tatcctttct ctaccttctc ctcagctttc tgagtattga atgttgttgt taactgtctc
tgaatacgat tgattttgtg tcatttttct cgtgcttcgt atttgttgat tgattgattt Δ5
acatgtcatg atttgaaatc tttgtattgc taattgctat tgctctgttt actgattttc Δ6
ttttctgggg ttatcaggtt tcgtgaaaag aatctatctt gagca... **ΔR**

D)

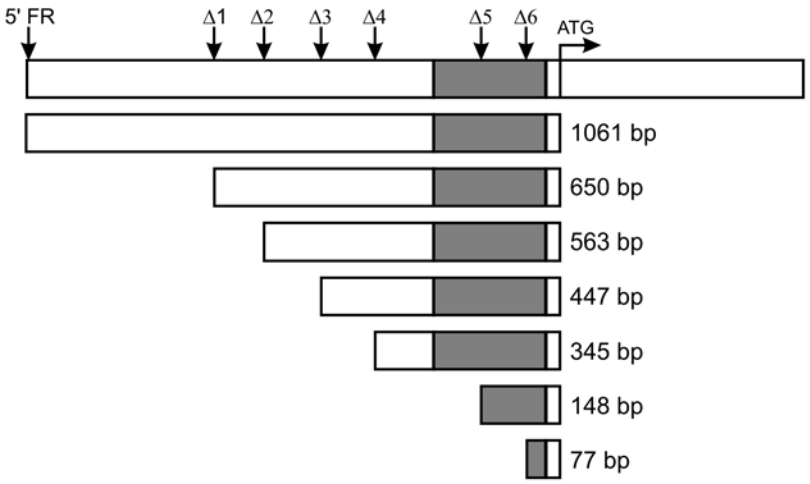


Figure 4.4. *RPL23aA* and *B* 5' deletion series. A) *RPL23aA* full 5' flanking region upstream of ATG start codon (dots indicate ORF). 5' leader intron, **bold** type. Reverse primer highlighted in yellow, forward primers highlighted pink, primer names in right hand margin (5' FR, 5' flanking region; Δ, deletion fragments; ΔR, reverse primer). B) Schematic of *RPL23aA* deletion fragments amplified from genomic sequence. Arrows indicate 5' end of each fragment, bent right angle arrow indicates start of ORF, grey shading indicates upstream intron sequence (included in fragments). Fragment sizes are listed in bp. C) *RPL23aB* full 5' flanking region, annotation as in A. D) Schematic of *RPL23aA* deletion fragments, annotation as in B.

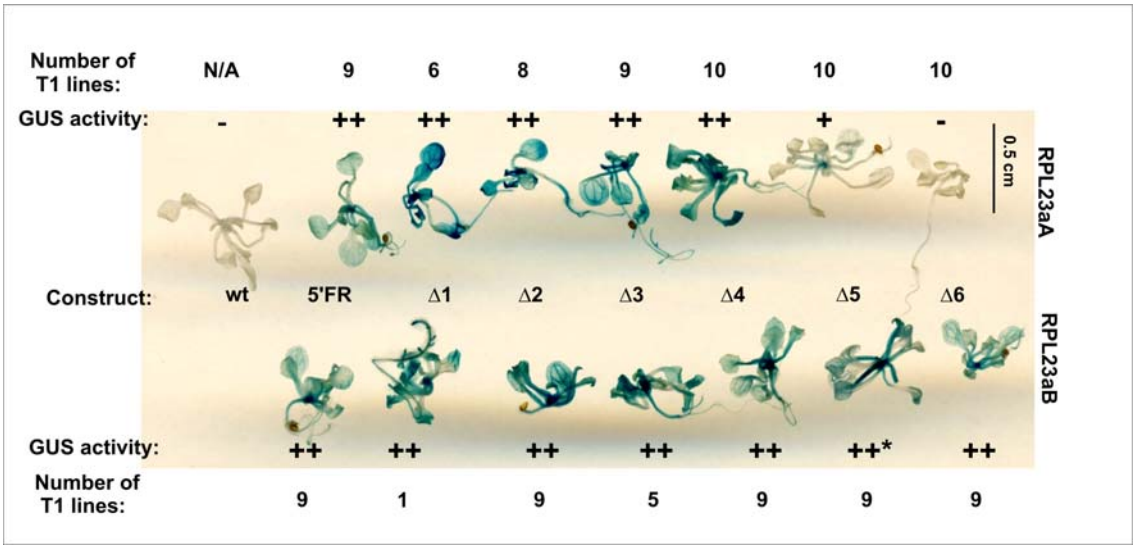


Figure 4.5. GUS activity in 16-19 day old wild type (wt) and *RPL23a* 5' RR T_2 seedlings. Seedlings shown are representative of staining patterns exhibited by all seedlings sampled for each construct. ++, positive for GUS activity, most or all of each seedling stained; +, weak positive for GUS staining, only small portion of each seedling stained, staining light; -, negative for GUS activity. All T_2 seedlings positive (++) for GUS activity except *RPL23aA* and *B* Δ5 (+, weak positive; ++*, not all seedlings in sample stained in ++ pattern), and RPL23aAΔ6 seedlings (-, negative). Number of independent T_1 lines used to generate T_2 plants shown above (*RPL23aA*) and below (*RPL23aB*) seedlings.

each parental line were planted in soil and grown to maturity (~7-8 weeks old), at which point the following series of 11 tissues was collected and assayed for GUS activity: basal rosette leaves, bracts, stems, unopened buds, buds at 0 days post-anthesis (0 DPA, tips of petals just emerging from bud), open flowers, elongating carpels/siliques less than 6 mm, siliques 6-10 mm, siliques >10 mm, drying/yellowing mature siliques, and root. Tissues from wild type control plants were collected at the same stages as all transgenic tissues and the control tissues were all negative for GUS activity.

Results of the GUS assays of mature plant tissues, vegetative and floral, are summarized in Tables 4.3 and 4.4. While the GUS expression patterns driven by the *RPL23a* regulatory fragments were much more complex in different tissues of mature Arabidopsis than in seedlings, some trends did emerge. In general, the full-length 5' RRs from both *RPL23aA* and *RPL23aB* conferred *GUS* expression in mitotically and developmentally active tissues such as root (especially in root vasculature and lateral root primordia; Figure 4.6 A-C), leaf margins and vasculature (Figure 4.6 D), developing tissues of the androecium and gynoecium, and elongating carpels (Figure 4.7). GUS activity was also seen at cut sites where tissue was removed from plants, especially at the ends of sections of stem and leaf, and at other wound sites where leaves and bracts were torn, folded, or otherwise mechanically damaged before or during harvest (Figure 4.6 E, F).

In keeping with higher levels of expression in active tissues, GUS activity was weak or absent in the fully mature tissues of leaf, bract (Table 4.3), and drying siliques (Table 4.4) from *RPL23aA* and *B* 5'FR T_2 plants. Siliques of transgenic plants carrying every construct except the *RPL23aA*Δ6 deletion showed reduced GUS expression during maturation. Staining was seen along the length of elongating carpels/siliques and in developing seed, but was reduced to staining at terminal sites (corresponding to the former stigmatic and floral abscission zone regions) and seed pods in mature siliques (Figure 4.7). Maturing/yellowing siliques in all *RPL23aB* T_2 lines and in most (Δ3-Δ6) *RPL23aA* lines showed staining only of seed pods, and by the time of dehiscence in most fully mature (dried) siliques, no GUS activity was noted (Figure 4.7). GUS activity was retained primarily in the replum (septum between the valves of the silique) and funiculi of siliques during maturation.

Construct	Root	Leaf	Bract	Stem	Construct	Root	Leaf	Bract	Stem
Wild type	-	-	-	-					
A5'FR	++	-	-	$++^{c}$	B5'FR	++	-	$+^{e}$	$++^{c}$
AΔ1	++	++	$+^{c,e}$	$++^{c}$	BΔ1	++	$+^{e}$	$+^{e}$	$++^{c}$
AΔ2	++	$++^{v,c}$	$+^{c,e}$	$++^{c}$	BΔ2	++	+	$+^{e}$	$++^{c}$
AΔ3	++	$+^{v,e}$	$+^{v,c}$	$++^{c}$	BΔ3	++	-	$+^{v,c}$	$++^{c}$
AΔ4	++	$+^{v,c,e}$	$+^{v,c,e}$	$++^{c}$	BΔ4	++	-	$+^{c}$	$+^{c,e}$
AΔ5	+	-	-	$+^{c}$	BΔ5	++	$+^{v}$	$+^{v,c}$	$++^{c,e}$
AΔ6	-	-	-	-	BΔ6	++	-	$+^{c,e}$	$++^{c,e}$

Table 4.3. GUS activity in wild type and *RPL23aA* and *B* 5' RR T_2 plants, vegetative tissues. ++, positive for GUS activity, most or all tissues in sample are stained/stain is dark; +, weak positive for GUS staining, only some tissue in sample is stained/staining light; -, negative for GUS activity. Superscripts indicate specific regions of concentrated staining where applicable: c, cut sites/ends of stems; e, end(s)/margins (leaf/bract); v, venation.

Construct	Unopened Bud						0 DPA Bud						Open Flower						Siliques			
			stamens			carpel			stamens			carpel			stamens			carpel					
	sepal	petal	anth	fil't	stigma	ovary	sepal	petal	anth	fil't	stigma	ovary	sepal	petal	anth	fil't	stigma	ovary	<6 mm	6-10mm	>10 mm	mat.
Wildtype	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A5'FR	-	+^{v}	++	+	++	++	-	+^{v}	++	+	+^{sp}	++	-	-	++^{p}	-	-	++^{r,s}	+^{s}	+^{e}	+	-
AΔ1	++	++	++	+	++	++	-	+^{v}	++	++	++	++	-	-	++^{p}	++	++	++^{r,s}	++	++	++	+
AΔ2	++	+^{v}	++	+	++^{sp}	++	++	+^{v}	++	++	++	++	+^{v}	+^{v}	++^{p}	++	++	++^{r,s}	++	++	++	+
AΔ3	++^{v}	+^{v}	++	++	++	++	++^{v}	+^{v}	++	+	++	+	+^{v}	+^{v}	+^{p}	++	++	+	++^{e}	++^{e}	+^{e}	++^{e}
AΔ4	++^{v}	++^{v}	++	+	+	+	+^{v}	+^{v}	++	+	++	++	+^{v}	+^{v}	++^{p}	+	++	++	+	+^{e}	+^{e}	+^{e}
AΔ5	-	-	+	-	+	+	-	+	+	-	+^{sp}	+	-	-	+^{p}	-	-	+	+	+	-	+
AΔ6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
B5'FR	+^{v}	+^{v}	++	+	++	++	+^{v}	+^{v}	++	++	++	++	+^{v}	+^{v}	++^{p}	++	++	++^{r}	++	++	+^{e}	+^{e}
BΔ1	++	+	++	++	++^{sp}	++	++	++	++	++	++^{sp}	++	+^{v}	+^{v}	++^{p}	++	++	++^{r}	++	++^{e}	n/a	++^{e}
BΔ2	++^{v}	+^{v}	++	++	++^{sp}	++	+^{v}	+^{v}	++	+	+^{sp}	++	+^{v}	+^{v}	++^{p}	++	++	++^{r}	++	++^{e}	++	++
BΔ3	++^{v}	+^{v}	++	+	++^{sp}	++	++^{v}	+^{v}	+	++	+	++	+^{v}	+^{v}	++^{p}	++	++	++^{r}	++	++^{e}	+^{e}	+
BΔ4	+^{v}	+^{v}	-	+	+	++^{r}	+^{v}	+^{v}	-	+	+	+^{r}	+^{v}	+^{v}	-	+	+	+^{r}	+^{e}	+^{e}	+^{e}	+
BΔ5	+^{v}	+^{v}	-	++	+	++^{r}	+^{v}	+^{v}	-	+	+	+^{r}	+^{v}	+^{v}	-	+	+	+^{r}	++	++^{e}	+^{e}	+
BΔ6	++^{v}	++^{v}	-	+	+	++^{r}	+^{v}	+^{v}	-	+	+	+^{r}	+^{v}	+^{v}	-	+	+	+^{r}	++	++^{e}	+^{e}	+

Table 4.4. GUS activity in wild type and 5' RR T_2 plants, floral tissues. anth, anther; fil't, filament; mat, mature/drying siliques; ovary, style and ovary. ++, positive for GUS activity, most or all tissues in sample stained/stain is dark; +, weak positive for GUS staining, only some tissue in sample is stained/staining light; -, negative for GUS activity; n/a, tissue not available. Superscripts indicate specific regions of concentrated staining where applicable: e, ends of siliques; p, pollen; r, receptacle; s, seed; sp, stigmatic papillae; v, venation.

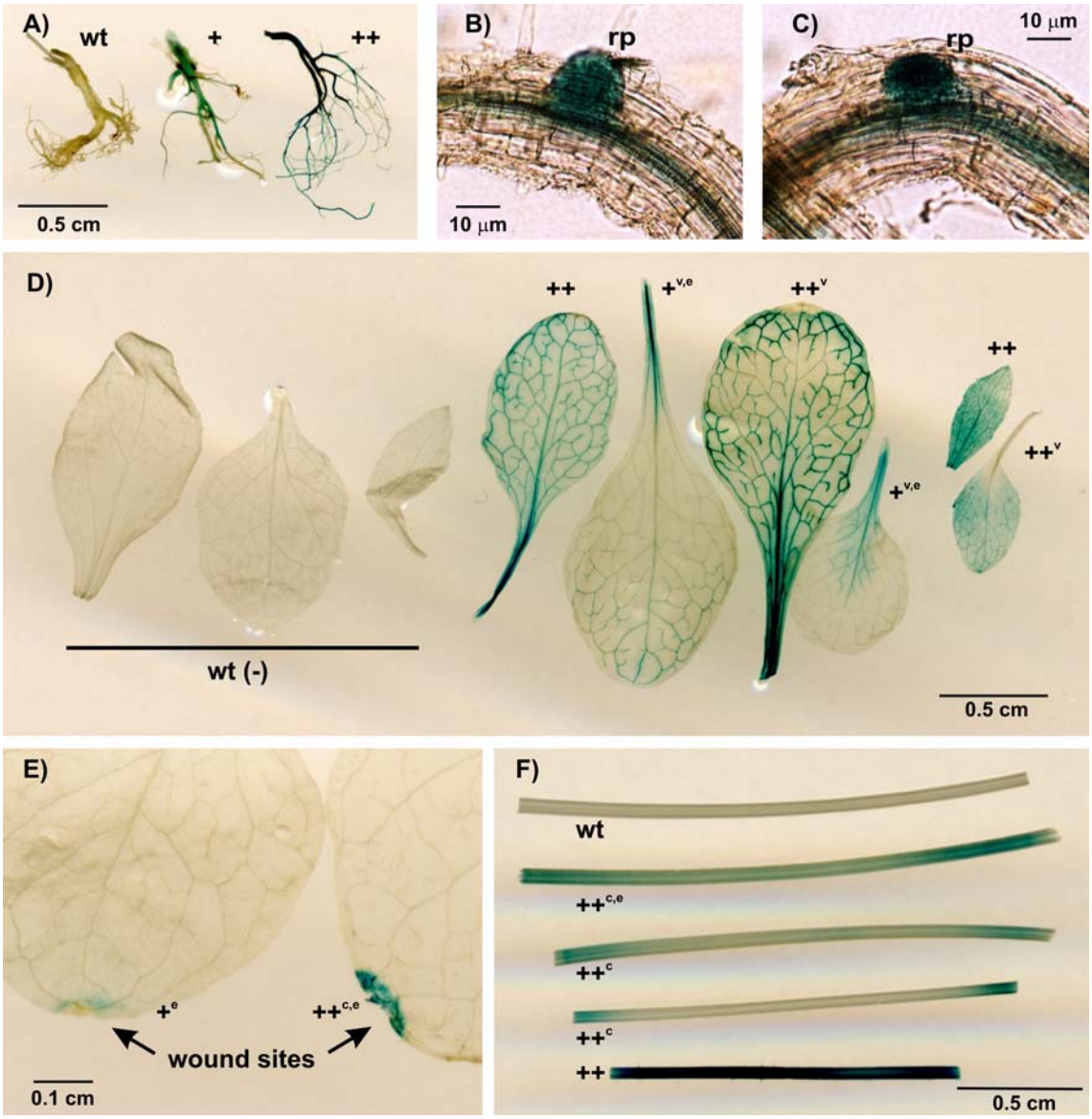


Figure 4.6. GUS activity in wild type (wt) and representative samples from *RPL23aA* and *B* 5' RR T_2 plants, vegetative tissues. ++, positive for GUS activity; +, weak positive for GUS staining; -, negative for GUS activity. Superscripts indicate regions of concentrated staining where applicable: c, cut sites/ends of stems; e, end(s)/margins (leaf/bract); v, venation. A) Portions of root removed from mature wild type and T_2 plants, middle (+) root from *RPL23aAΔ5* T_2. B) and C) T_2 roots showing lateral root primordia, rp. D) Sample of leaves and bracts from wild type and *RPL23aA* and *B* transgenic plants. E) Margins of mature T_2 leaf (left) and bract (right) showing GUS activity at wound sites. F) Sections of stem from wild type and T_2 plants.

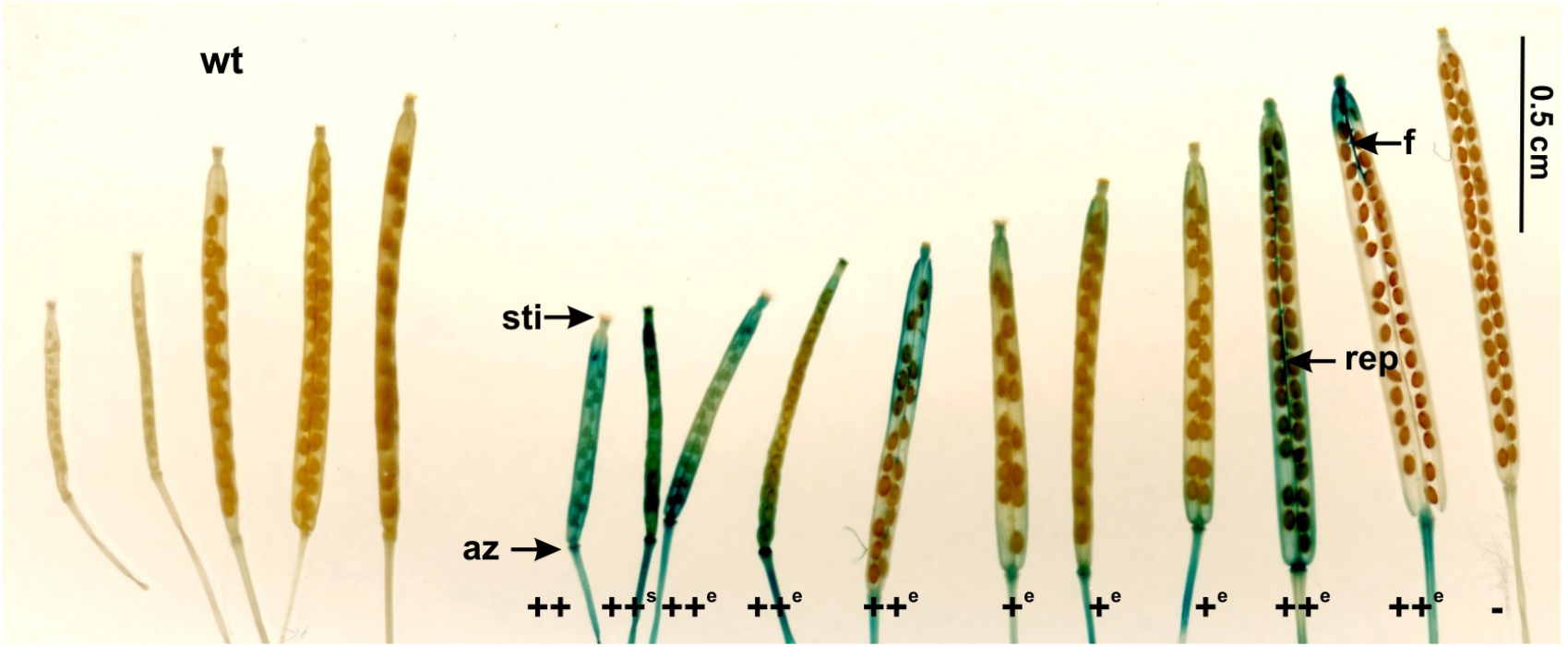


Figure 4.7. GUS activity in wild type (wt) and representative samples from *RPL23aA* and *B* 5' RR T_2 plants, developing/elongating siliques. Siliques shown correspond to categories listed in Table 4.4, arranged in order, <6 mm, 6-10 mm, >10 mm, and mature/drying siliques. az, floral abscission zone; f, funiculus; rep, replum; sti, stigmatic region. ++, positive for GUS activity; +, weak positive for GUS staining; -, negative for GUS activity. Superscripts indicate specific regions of concentrated staining where applicable: e, ends of siliques; s, seed.

Despite an overall similarity in GUS activity conferred by both *RPL23a* full-length 5'FRs, there were some notable differences between genes. While *RPL23aB*5'FR plants showed some degree of GUS activity in all floral tissues, *RPL23aA*5'FR transgenics consistently lacked GUS expression in sepals, and showed weak expression in petals. Likewise, no GUS expression was seen in the leaves or bracts from *RPL23aA*5'FR plants while *RPL23aB*5'FR plants showed staining at the ends and leaf margins of bracts. Both *RPL23a* 5'FRs directed GUS activity in the reproductive organs (stamens and carpels) of buds and flowers but the staining of carpels in *RPL23aB*5'FR plants was more persistent as carpels elongated into siliques following anthesis. The persistence of GUS activity in siliques was seen in all of the *RPL23aB* Δ T_2 plants but only in some (Δ1-Δ4) of the *RPL23aA* transgenics. Differences in expression between the 5'RR deletion fragments from the two *RPL23a* genes were also observed in stamens. Anthers (and carpels) were the last tissues to lose GUS activity in *RPL23aA* 5' deletion transgenics, persisting in *RPL23aA*Δ5 plants with only 145 bp of 5' RR. In contrast, GUS activity was not detected in the anthers of plants carrying the last three *RPL23aB* 5' deletions, with up to 354 bp of upstream RR (Table 4.4, Figures 4.8-4.9). The inverse staining pattern was seen in filaments; while GUS activity in the filaments of stamens from *RPL23aA*Δ4 and Δ5 T_2s was weak or absent (while anthers were stained), staining of filaments and connective tissue in anthers in *RPL23aB* transgenics persisted even when the sporogenous tissues of their anthers did not show GUS activity. All *RPL23aA* and *B* transgenics that did show staining in anthers clearly expressed GUS in the sporogenous tissue of anthers rather than in connective tissue; GUS activity became restricted to pollen as buds opened and mature flowers developed (Figure 4.9).

A comparison of different deletion lines for *RPL23aA* and *B* was utilized to determine important upstream RRs for each gene. Little difference in expression pattern was seen between transgenic plants carrying the *RPL23aA*Δ1, Δ2, Δ3, or Δ4 constructs, which contain between 887 and 295 bp of 5' RR upstream of the ATG start codon. In contrast to the *RPL23aA*5'FR plants, which showed low or no GUS expression in many tissues, the majority of the *RPL23aA*Δ1-Δ4 T_2 plants showed relatively high amounts of GUS activity in most vegetative (leaf, bract, stem) and floral (sepals, petals, mature siliques) tissues examined. As in seedlings, *RPL23aA*Δ5 plants (carrying 145 bp of 5'RR

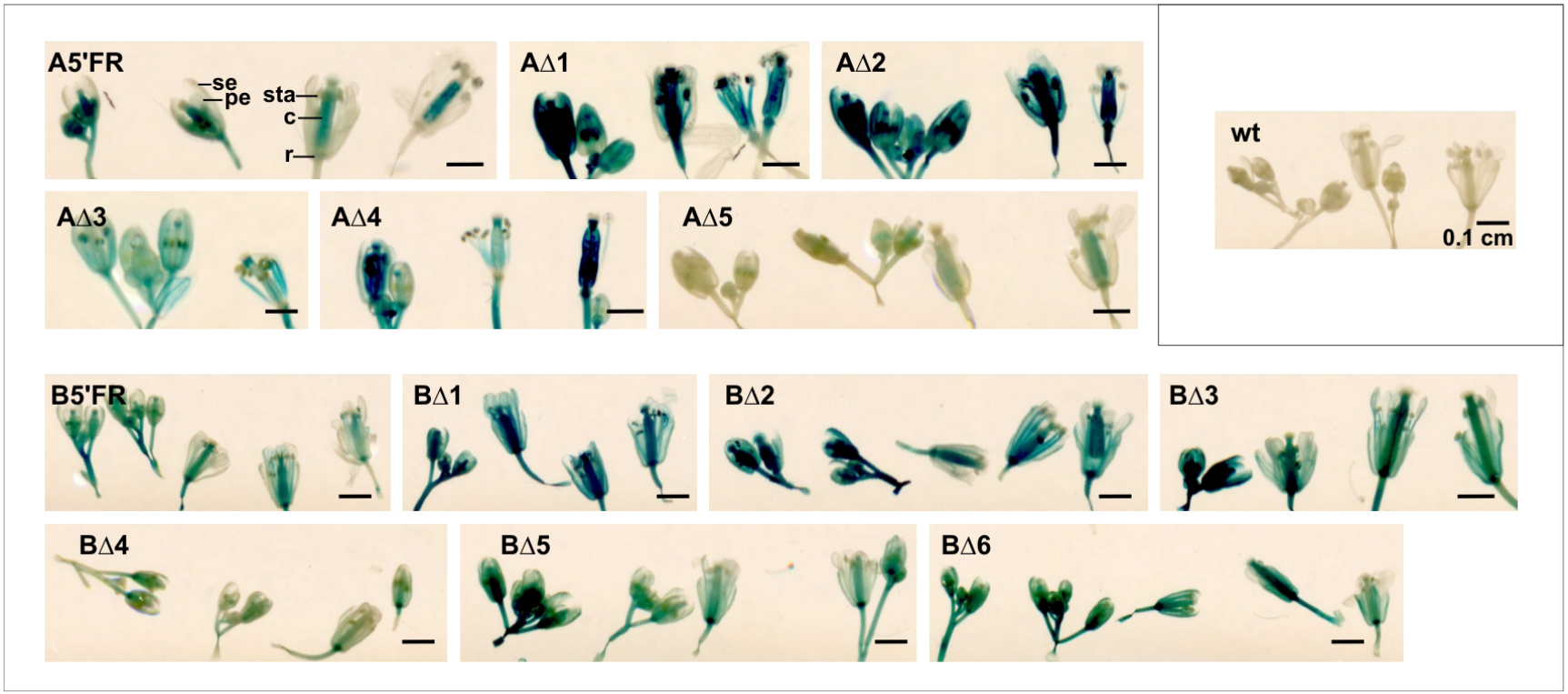


Figure 4.8. GUS activity in wild type (wt) and *RPL23aA* and *B* 5' RR T_2 plants, buds and flowers, one panel per construct. Each panel represents a selection of unopened buds, 0 DPA buds, and open flowers; staining categories are listed in Table 4.4. *RPL23aA*? 6 transgenics (not shown) are unstained, identical to wild type. All scale bars represent 0.1 cm. c, carpel; pe, petal; r, receptacle; se, sepal; sta, stamen.

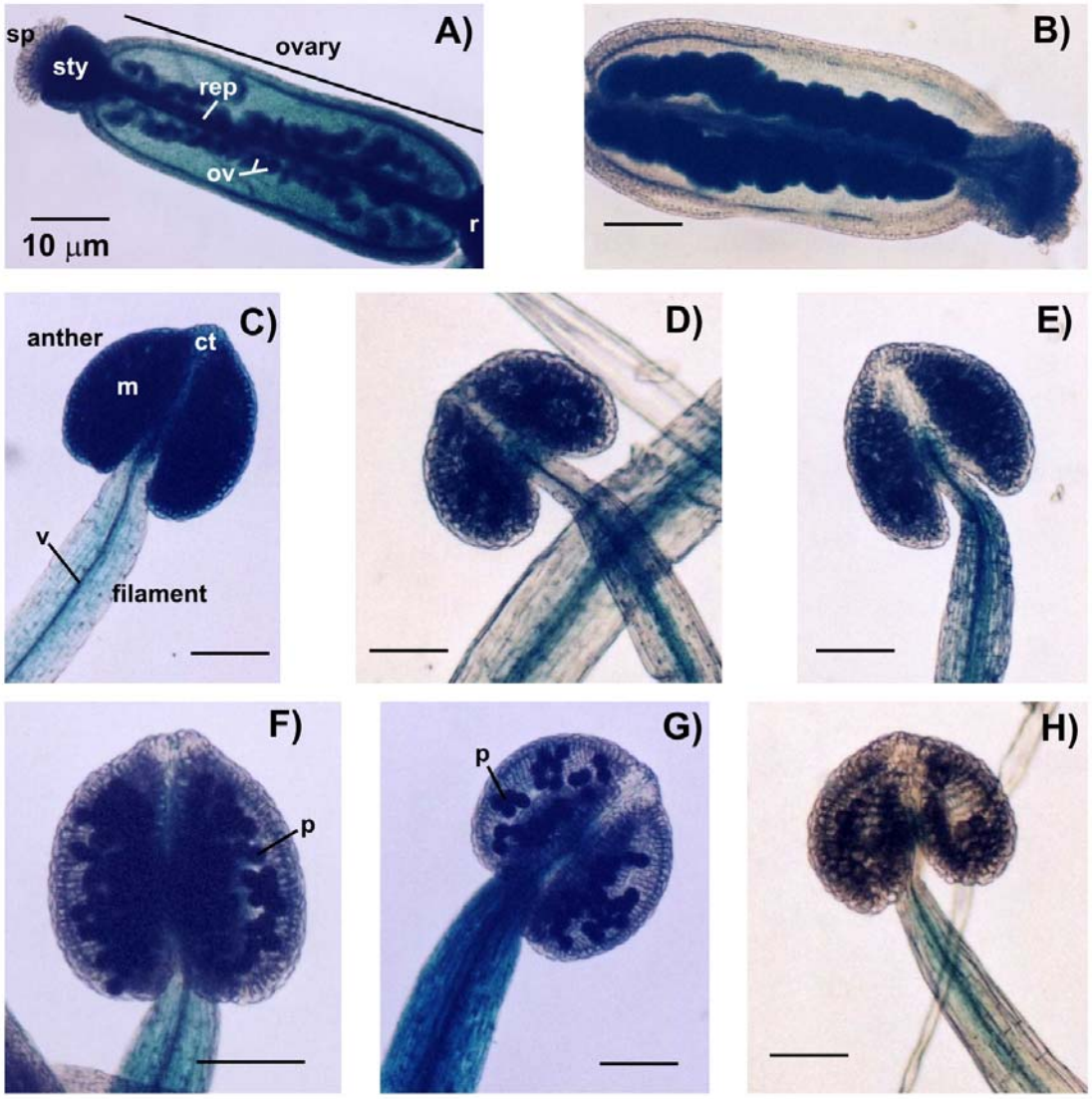


Figure 4.9. GUS activity in *RPL23aA* and *B* 5' RR T_2 plants. All scale bars represent 10 μm. A-B) Carpels showing whole ovary, all other floral organs removed. sp, stigmatic papillae; sty, style; rep, replum; ov, ovules; r, receptacle. C-I) Isolated stamens. m, microsporangium; ct, connective tissue; v, vascular strand; p, pollen. C-E) Stamens showing GUS activity in anthers during pollen development. F-G) Stamens showing GUS activity in pollen. H-I) Stamens from *RPL23aB*Δ4-Δ6 plants showing GUS activity in filaments but not in anthers.

upstream of the *RPL23aA* ORF) showed greatly reduced GUS expression compared to the other *RPL23aA* transgenics, showing weak expression in root and stem and no expression in leaves, bracts, or any sterile (non-gametogenic) floral tissues (sepals, petals, filaments). Any GUS activity in *RPL23aA*Δ5 plants was shown primarily as weak expression in anthers, carpels, and elongating siliques. *RPL23aA*Δ6 (66 bp of 5' RR) failed to direct GUS expression in any vegetative or floral tissues.

Transgenics carrying *RPL23aB* deletion constructs showed few broad differences in terms of GUS activity in most tissues, and differences in expression between *RPL23aB* seedlings were not entirely indicative of the differences between tissue types in flowering plants. While T_2 seedlings demonstrated a reduced GUS expression only in *RPL23aB*Δ5 transgenics, tissues of mature Arabidopsis T_2s showed a more complex pattern. GUS expression in vegetative tissues did not differ significantly between the *RPL23aB* deletion plants and plants carrying the full 5'FR, although while no GUS staining was observed in mature leaves of B5'FR plants, *RPL23aB*Δ1, Δ2, and Δ5 transgenics all showed some degree of staining in leaves (Table 4.3). Reduced GUS activity with serial 5' deletions of the *RPL23aB* upstream RR was only apparent in floral tissues; *RPL23aB*Δ4, Δ5, and Δ6 plants (carrying 345-77 bp of 5'FR) showed a consistent reduction of GUS staining in the stamens and carpels of buds and flowers. T_2 plants carrying the *RPL23aB*Δ4, Δ5, and Δ6 deletions showed reduced GUS activity in carpels (stigma, style, and ovary), and expression in anthers was completely abolished, although GUS activity was retained in staminal filaments (Figure 4.9H, I). In addition to the sterile tissues of the androecium, *RPL23aB*Δ4, Δ5, and Δ6 plants showed strong GUS activity in the receptacle regions of buds and flowers (Table 4.4, Figure 4.8). GUS staining in the receptacles of *RPL23aB5*'FR-Δ3 transgenics did not appear strongly as a discrete region of activity until flowers were open and carpels began to elongate.

4.3.3. RT-PCR confirmation of transcription in transgenic plants

The histochemical GUS assay detects the presence or absence of a functional GUS protein, so it was unknown whether the decreased and absent GUS activities in the *RPL23aA*Δ5 and Δ6 plants was due to transcriptional repression or inhibition of translation. RT-PCR was used to determine the absence or presence of *RPL23aA*Δ5- and, more importantly, *RPL23aA*Δ6-regulated *GUS* gene transcripts. The same

technique was also used to determine if *RPL23aA* and *B* transcripts from deletion constructs containing complete or partial upstream introns were processed. Primers were designed to amplify the 5' RRs of the transgenic *RPL23aΔ::GUS* transcripts; a *GUS*-specific reverse primer (Figure 4.10) was used for reverse transcription, followed by PCR amplification using the *GUS* primer and gene-specific primers for each construct (see Materials and Methods). T_2 seedlings from independent *RPL23aA* and *B* 5'FR, Δ4, Δ5, and Δ6 lines (three lines tested per construct) were used to examine *GUS* transcripts. The 5'FR constructs for each gene were chosen as a full-length comparison that included the entire upstream intron, the Δ4 constructs represent the shortest fragments that contained the entire upstream intron with a sizable portion of sequence 5' to the intron, and the Δ5 and Δ6 constructs were the primary focus of the RT-PCR study and contained the full upstream intron sequence with very little 5' border (*RPL23aA*Δ5) or partial upstream intron (Δ6) sequence.

RT-PCR amplification of transcripts from *RPL23aA* transgenics yielded predicted fragment sizes based on mapped transcript start sites (included in 5'FR and Δ4 constructs) and splicing of the upstream intron in 5'FR, Δ4, and Δ5 T_2 plants (Table 4.5, Figure 4.11). The splicing of the upstream intron from the *RPL23aA*5'FR, Δ4, and Δ5::*GUS* transcripts was confirmed via sequencing of the RT-PCR fragments using the pC-GUS-R2 reverse primer. Interestingly, while the 5' leader intron of the *RPL23aA* 5'FR and Δ4 transcripts was spliced at the canonical $^{5'}$GT-AG$^{3'}$ splice sites seen in the wild type transcripts characterized by 5' RACE (see above), the Δ5 transcript was spliced at either $^{5'}$TT-AT$^{3'}$ or $^{5'}$CT-GT$^{3'}$ sites just downstream of the wild type leader intron splice sites (Figure 4.12). The *RPL23aA*Δ6 transcript showed no splicing of the upstream intron and included the entire unspliced Δ6 RR fragment in its 5' UTR. *GUS* transcripts were thus confirmed in the RPL23aAΔ5 and Δ6 transgenics, indicating inefficient and repressed translation, respectively, rather than inhibition of transcription of the reporter gene construct in these plant lines.

RT-PCR products from *RPL23aB* transgenics showed reporter gene transcripts in all lines, but multiple bands were amplified from plants carrying the Δ5 and Δ6 deletion constructs. While RT-PCR yielded fragments of the expected size for each construct

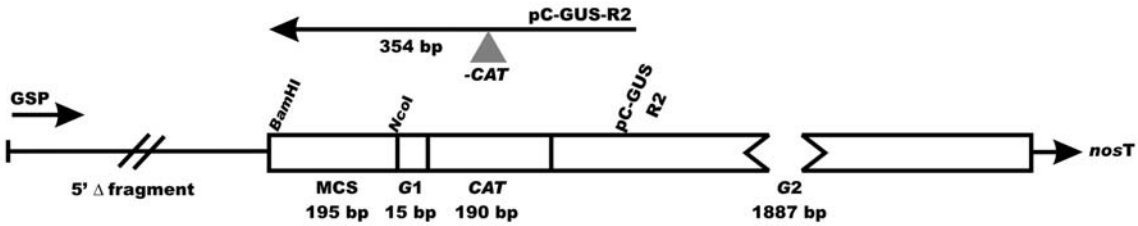


Figure 4.10. *GUS* reverse primer (pC-GUS-R2) and gene-specific forward primer (GSP) used for reverse transcription and PCR of deletion fragment/reporter transgene transcripts, shown positioned on portion of pCAMBIA1381Z T-DNA. MCS, multiple cloning site sequence 3' to *RPL23a* 5' RR fragment; *G*1, *GUS* first exon; *CAT*, *CATALASE* intron; *G*2, *GUS* second exon; *nos*T, *nopaline synthase* terminator. Left-facing arrow above T-DNA indicates length between pC-GUS-R2 primer and *Bam*HI site, including portion of MCS (deleted *CATALASE* intron shown as **grey** triangle).

Construct	Forward primer	Expected fragment size
*RPL23aA*5'FR	L23aA-RTintF	401 bp
*RPL23aA*Δ4	L23aA-RTintF	401 bp
*RPL23aA*Δ5	L23aAΔ5 deletion primer	390 bp
*RPL23aA*Δ6	L23aAΔ6 deletion primer	420 bp
*RPL23aB*5'FR	L23aB-RTintF2	395 bp
*RPL23aB*Δ4	L23aB-RTintF2	395 bp
*RPL23aB*Δ5	L23aBΔ5 deletion primer	502 bp
*RPL23aB*Δ6	L23aBΔ6 deletion primer	431 bp

Table 4.5. Expected RT-PCR amplification product sizes for transgenic *RPL23a*Δ plant lines. Each template was subjected to RT-PCR with a common GUS-specific reverse primer (pC-GUS-R2) and specific forward primer as listed. Expected fragment size is based on splicing of *CATALASE* intron from *GUS*, splicing of upstream intron in 5'FR, Δ4, and *RPL23aA*Δ5 plants, and most common 5' RACE-determined transcription initiation sites for 5'FR and Δ4 constructs.

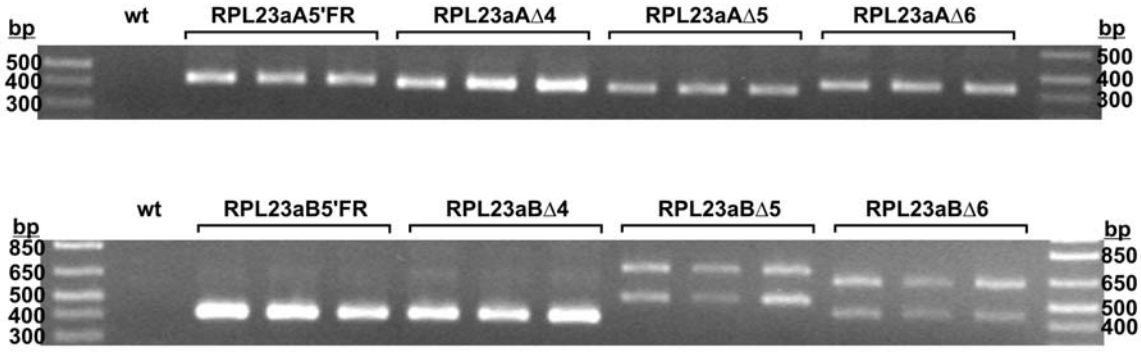


Figure 4.11. Ethidium bromide-stained gels of RT-PCR products (*GUS*::*RPL23a* 5' RR transcript fragments) amplified from transgenic *RPL23aA* and *B* deletion series T_2 seedlings. Reverse transcription from total RNA was performed using a *GUS* ORF-specific reverse primer, followed by PCR using *GUS* reverse and *RPL23aA* and *B* 5' RR fragment-specific forward primers. RT-PCR using total RNA from non-transgenic (wild type, wt) plants was included as a negative control. Three different transgenic lines were tested per construct, shown as three lanes under brackets for each construct.

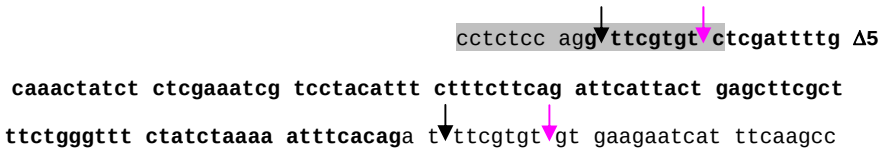


Figure 4.12. Entire 5' RR deletion fragment cloned upstream of *GUS* ORF in *RPL23aA*Δ5 transgenics, showing features of transcript amplified via RT-PCR. *RPL23aA*Δ5 primer highlighted in grey, upstream intron with canonical splice sites shown in bold type, two sets of arrows (black and **pink**) indicate alternate transcript splice sites determined in *RPL23aA*Δ5 T_2 plants.

(Table 4.5; Figure 4.11, bottom bands), there were also higher molecular weight fragments amplified from *RPL23aB*Δ5 and Δ6 transgenics (Figure 4.11, top bands in *RPL23aB*Δ5 and Δ6 lanes). Sequencing of *RPL23aB5*'FR and Δ4 transcript fragments confirmed splicing of the upstream intron at the $^{5'}$TC-TA$^{3'}$ splice sites most commonly identified in 5' RACE products from wild type plants (see above). As seen in *RPL23aA* transgenics, no splicing of partial upstream intron sequences was seen in *RPL23aB* deletion series plants; sequencing of cloned transcripts confirmed the presence of the entire *RPL23aB*Δ5 and Δ6 deletion fragments in the 5' UTRs of their respective transcripts. In order to check whether the higher molecular weight bands obtained from the *RPL23aB*Δ5 and Δ6 plants were artifacts of the primers used for RT-PCR, a number of control experiments were performed. Although all templates were DNase I-treated, reactions excluding the SuperScript II reverse transcriptase, yielding no product, confirmed that there was no contamination of the RNA template with DNA. RT-PCR using single primers (GUS primer or *RPL23a*-specific primer alone) also confirmed that the higher molecular weight bands were not single-primer artifacts. Numerous efforts to sequence transcripts from *RPL23aB*Δ5 and Δ6 plants with both *GUS*- and *RPL23aB*-specific primers gave poor quality results, but the presence of the deletion fragments in the 5' UTRs could be confirmed, as mentioned above. Unexpectedly, the higher molecular weight products found in *RPL23aB*Δ5 and Δ6 plants appeared to be the result of a failure in the splicing of the *CATALASE* intron present between the *GUS* first and second introns. The retention of the 190 bp *CATALASE* intron resulted in a 692 bp fragment of the *RPL23aB*Δ5::*GUS* transcript and a 621 bp fragment of the Δ6::*GUS* transcript, corresponding exactly to the higher molecular weight bands seen for the transgenics.

4.4. Discussion

In order to determine some of the regulatory mechanisms controlling multiple r-protein gene family members in plants, I have conducted an investigation into the upstream RRs of the two genes encoding RPL23a in Arabidopsis. My previous study of relative transcript abundances for both genes in a variety of wild type Arabidopsis tissues showed that while both *RPL23aA* and *B* transcripts were ubiquitous, there were differences in transcript levels for each gene between tissues, and between the two genes

(McIntosh and Bonham-Smith, 2005; Chapter 3, this volume). The quantitative differences in regulation of the two *RPL23a* genes was not surprising given the relatively low sequence identity between the two upstream RRs (McIntosh and Bonham-Smith, 2005; Chapter 3, this volume), but what was interesting was the amount of qualitative similarity between *RPL23aA* and *B* in terms of overall expression patterns conferred by their different upstream regions. In accordance with the earlier *RPL23aA* and *B* transcript abundance study, in this study the full 5' RRs for both genes conferred reporter gene expression most strongly in mitotically and developmentally active tissues. GUS activity in plants carrying the complete 5' flanking region for each *RPL23a* gene was strongest in seedlings (Figure 4.5), root (especially root meristems; Table 4.3, Figure 4.6A-C), and gametogenic tissues (anthers and developing carpels; Table 4.4, Figures 4.8-4.9). This specificity of GUS expression for mitotically active tissues is in agreement with expression patterns observed for other r-protein genes in Arabidopsis (Van Lijsebettens *et al*., 1994; Williams and Sussex, 1995), tobacco (*Nicotiana tabacum*; Marty and Meyer, 1992; Gao *et al*., 1994; Dai *et al*., 1996), maize (Larkin *et al*. 1989; Lebrun and Freyssinet, 1991; Joanin *et al*., 1993; Chevalier *et al*., 1996), canola (*Brassica napus*; Bonham-Smith *et al*., 1992), pea (*Pisum sativum*; Strafstrom and Sussex, 1992; Moran, 2000), petunia (*Petunia hybrida*; Lee *et al*., 1999), and potato (*Solanum tuberosum*; Taylor *et al*., 1992). In particular, the upstream regulatory sequence from Arabidopsis *RPL16A* (*RPL11A* according to the nomenclature of Barakat et al., 2001) was found to direct reporter gene expression behind the root meristem, in lateral root primordia, the stele, and in developing anthers and pollen (Williams and Sussex, 1995), patterns also seen in *RPL23aA* and *B* 5'FR transgenics (Figures 4.6A-C, 4.9).

Unlike the somewhat ambiguous changes in *RPL23aA* and *B* transcript abundance following previous wounding experiments (McIntosh and Bonham-Smith, 2005; Chapter 3, this volume), the 5' RRs of both *RPL23a* genes clearly induced reporter gene expression at wound sites (Figure 4.6E-F). The strong induction of expression at the cut sites in stem tissues (Figure 4.6F) could also explain the high levels of *RPL23aA* and *B* transcripts detected in stem during the previous expression study (McIntosh and Bonham-Smith, 2005; Chapter 3, this volume). Although most of each

RPL23aA and *B*5'FR T_2 stem section did not show GUS activity, the ends of the stems, following from the cut sites where the tissue was harvested, show high levels of GUS expression that were induced immediately upon wounding before tissues were fixed in acetone. This induction of expression at wound sites is in agreement with *L25* expression in tobacco, which showed a 3-fold increase in leaves following wounding (Gao et al., 1994). While neither *RPL23aA* or *B*5'FR plants showed overall expression in mature leaf or bract tissues (Table 4.3), the *RPL23aB* full 5'FR conferred GUS expression at the cut petiole ends of mature bracts, showing greater wound inducibility across tissues than the *RPL23aA* 5'FR.

In addition to the differences between *RPL23aA* and *B*5'FRs in terms of wound induction, a lack of reporter expression in sepals and weak staining in petals of *RPL23aA*5'FR T_2 plants compared to *RPL23aB*5'FR transgenics was also identified (Table 4.4, Figure 4.8). Clearly the 5' RR of *RPL23aB*, even the 77 bp Δ6 fragment, confers more extensive expression in vegetative and sterile floral tissues than that of *RPL23aA*. While staining in sepals, petals, filaments, receptacles, and seed pods was much more persistent in all *RPL23aB* transgenics than in *RPL23aA* plants, GUS expression was abolished in anthers and was weak in developing ovules and seeds in plants carrying 345 bp or less of 5' RR (Δ4-Δ6 plants; Table 4.4, Figure 4.9H, I).

A dissection of the 5' RRs for both *RPL23a* genes via deletion fragments cloned upstream of GUS and expressed in stably transformed plants yielded surprisingly different results for the two genes. *RPL23aA*-driven GUS expression was almost abolished by the removal of 150 bp between 295 and 145 bp upstream of the ATG (region between Δ4 and Δ5), a region that includes a putative TATA box and both putative *telo* boxes of the *RPL23aA* 5' FR (Figure 4.13). *RPL23aA*Δ5 seedlings showed substantially reduced GUS activity compared to 5'FR-Δ4 seedlings (Figure 4.5), and mature Δ5 plants showed dramatically decreased staining in vegetative and floral tissues, reduced to weak staining only in roots, stem cut sites, anthers and carpels (Tables 4.3, 4.4; Figures 4.6, 4.8). A further 79 bp deletion, resulting in the Δ6 fragment of only 66 bp of *RPL23aA* 5' RR, completely abolished GUS activity in all seedlings and mature plant tissues (Tables 4.3, 4.4; Figure 4.5). Interestingly, while the first *RPL23aA* deletion (Δ1) transgenics, carrying an 887 bp fragment after deletion of 616 bp of the 5'FR,

A) ***RPL23aA***

```
1                                                                ggaga 5'FR
6     ggaggagcaa attgtttacc caaacaacca agacagttgt gtagtttact cggagatttc
66    ttctccttgt gttccggtga aacctggaaa tcatcgaaat ttaactgcca ggaacggtca
126   ccgttaccgt gaaactcgag cagcgagatc ccttcgtcgg aatctggact agctggctca
186   accatttttc cggcggcgga gagagacttt gtatttttt ttttttttt ttgagagaga
246   gagagagtgg taaggaaatg ggaaatttga agagacggcg agggtaaaat cggtaattaa
306   gaaagattga ggctagctac tacgactgtt gtggtagtaa gacattcctc cacagagata
366   cttggaccag cctttagaag agacgtctct gctaattatc cacgtggcgt tctctcaatg
426   gatttgggat ctcgcacgtg attggaggag tcatcataca cacgtgtaca gttgagattc
486   aactgcaatt aaacgacccc gtgacatgac atggttatga gtcacatatg tgccctctca
546   aactctatta tgtagtataa aaactgtatt agatacaacc gacccaagat atagaagaac
606   ctttgaaaaa tggtagaagc cagttcagct aaagctttct tcttcacaat tatcgtaaaa Δ1
666   tattagacat gtagagcagg cttgttgttc tcttattgtc aagcgtaagt aaatgattat
726   tgtttagttt cgtgaccaat tggtttgttt tttggttagc cattgttcga tttgctttgg
786   tttggtcaaa gattgtatgc tttcaagtca aaccaaagct gactaaattt tattctgtaa
846   taatttactt tggaatggca aaaccaaaga aaggtcatag acacaagaga agaagcaaat
906   atgtttctta cgccaaaaga cacgtttgta tggtaagaca aaagacacgt ttgtatgttt Δ2
966   cttacgccaa taccattacg ttaaaaataa actagtaact taaatttacg ctcactttct
1026  atttgggttt tacatgatga ttttggtcaa aggcctattt acaatgttgc aagcccatat
1086  tatcttttaa gtatcttttt gtgcaaccaa aagaatcagt gatcttatgc tcattttcat Δ3
1146  ttttttctaa tgaaattttt cgtcacttat ggcttatgcg tccgatggac gattatggta
1206  aaaggcccat ttattcaatc caaagcttac tgggttagat gatgattgtg gtaaaaaggc Δ4
1266  ccattataga cccaactaag tatcttaggg tttcagtttc cactataaaa cttgttgcgg
1326  cgtaagttag ggtttgaga atcagcggct tcacctctcc aggttcgtgt ctcgattttg Δ5
1386  caaactatct ctcgaaatcg tcctacattt ctttcttcag attcattact gagcttcgct Δ6
1446  ttctgggttt ctatctaaaa atttcacaga tttcgtgtgt gaagaatcat ttcaagcc ATG
```

B) *RPL23aB*

```
1       catgaa tttgagttag aggatggttg gaacaaaaaa acttagaagc tcgaatgacc 5'FR
57   ggtttttacc aaattctcat agaccatatt tgattctttt gatttacttc tggtgcagga
117  ctctctgtgc ttatggaagt tgatgttggg ggaaacaact ctcttgtaca gtggggaaaa
177  aacttcttct tcttctttct atcacatgaa aatcctcaag ggccattatt agtatgatca
237  gattataaaa ttgtaaggtt aggggcttta tgaggatttt gatggacttg ttacaatgtt
297  tacatataca ctcagcagca caatagattt ttgttaaact tacatgttat tcaagtaaaa
357  gtactatgta gatgttgaag tctaattgaa gaattagtta atgatagtct taaacacttg Δ1
417  attcacttgt catccaattt tggttttgcg catagtttct cttcttttat ttcctctcta
477  aaacaccaaa accaaacaaa atgtcatttt ccaatcctta aggtttcatt catttagtg Δ2
537  attttttggg tacaaaattg agcaatgtct agtgacgttt ttactcaaac tcataaacca
597  acattctaat cagaatcacg atttggactt tggtttgga ccttctttct acaccaactg Δ3
657  ggcttgtaat cagttggtag agaaagccca attatatatt atttaaggcc caaaataaat
717  cgatctaggg tttacggttt tgtttcattt tcacttccct tagaataaat aaaaaccacc Δ4
777  aagcaacttg gatctactct agggtttctg tttcgccgct caggttcgtc aatctctcga
837  tatcctttct ctaccttctc ctcagctttc tgagtattga atgttgttgt taactgtctc
897  tgaatacgat tgattttgtg tcatttttct cgtgcttcgt atttgttgat tgattgattt Δ5
957  acatgtcatg atttgaaatc tttgtattgc taattgctat tgctctgttt actgattttc Δ6
1017 ttttctgggg ttatcaggtt tcgtgaaaag aatctatctt gagcaATG
```

Figure 4.13. *RPL23aA* (A) and *B* (B) full 5' flanking regions as used in deletion series. Bold uppercase ATG indicates the start of the ORF. Upstream intron, **bold** type; repeat motifs found at 5' leader intron splice sites, underlined; black arrows, transcript start sites; TATA boxes highlighted in yellow; *telo* boxes (- orientaton; Lenvik et al., 1994; Trémousaygue et al., 1999), bold **blue** type; site II motifs (- orientation; Trémousaygue et al., 2003), bold **orange** type; pollen-specific motifs ($^{5'}$AGAAA$^{3'}$, Bate and Twell, 1998; $^{5'}$GTGA$^{3'}$, Rogers et al., 2001), bold **light blue** type; root-specific motifs ($^{5'}$ATATT$^{3'}$; Elmayan and Tepfer, 1995), bold **green** type. Deletion series forward primers highlighted pink, names in right hand margin (5' FR, 5' flanking region; Δ, deletion fragments).

showed almost the same expression pattern as the 5'FR plants, the Δ2 T_2s appeared to gain additional GUS activity. Unlike the rest of the *RPL23aA* transgenics, *RPL23aA*Δ2-Δ4 plants (555-295 bp 5' to ORF) showed GUS activity in the sepals of buds and flowers (Table 4.4, Figure 4.8), suggesting that an inhibitory element may exist in the 332 bp between the Δ1 and Δ2 fragments.

In contrast to the patterns seen in *RPL23aA* 5' deletion series plants, *RPL23aB* 5' deletion transgenics all showed GUS activity in seedlings and mature Arabidopsis tissues. Little difference was seen between *RPL23aB*5'FR-Δ6 transgenics in seedlings and vegetative tissues, except for reduced staining of Δ5 seedlings. Only the floral tissues of the *RPL23aB* plants showed any significant differences between 5' RR fragments; *RPL23aB*Δ4-Δ6 plants, carrying between 345 and 77 bp of 5' RR, did not show any GUS activity in the anthers or pollen of buds and flowers and showed greatly reduced staining in carpels (Table 4.4, Figure 4.9H, I). This suggests that the 102 bp region between the *RPL23aB*Δ3 and Δ4 deletion primers contains *cis*-elements required to direct expression in anthers and developing pollen; this region contains two *PCNA* site II motifs (upstream of the *telo* boxes included in Δ4 fragments) and a pollen-specific element (Figure 4.13), but deleting two other pollen-specific motifs present in the *RPL23aB* 5'FR did not abolish staining in the anthers of Δ3 plants, suggesting at least some redundancy of these elements, if functional. While the site II motif directs expression in actively dividing tissues (Trémousaygue et al., 2003), there is no indication that it is pollen-specific. Unlike the case in Δ4-Δ6 plants, GUS activity did not appear to differ significantly between transgenics carrying the *RPL23aB*Δ1-Δ3 fragments and those carrying the full 5'FR, even in floral organs (Tables 4.3, 4.4; Figure 4.8).

In addition to the analysis of 5' deletion series transgenic plants, further characterization of the *RPL23aA* and *B* upstream RRs was carried out via transcription start site mapping (via 5' RACE) and RT-PCR using template from deletion series transgenics. The mapping of transcription initiation sites in *RPL23aA* and *B* yielded some interesting results; both genes show multiple transcript start sites and possess introns in their 5' leader sequences (Figures 4.2-4.3). In contrast to mammalian r-protein

transcripts, which initiate at a C nucleotide within a polypyrimidine tract (Perry, 2005), primer extension analysis and comparison of GenBank cDNAs with genomic sequence determined that *RPL23aA* and *B* transcripts can initiate at a G, A, T, or C nucleotide, with A (*RPL23aB*) and T (*RPL23aA*) being the most common start sites. Although *RPL23aB* transcripts showed 5' UTRs that differed by as much as ~60 nt, RT-PCR amplification from wild type tissue failed to yield the longer transcript type, indicating that the shorter species of *RPL23aB* transcript is probably the most common. As in 60% of mammalian r-protein genes (Perry, 2005), a search of the *RPL23aA* and *B* genomic sequences showed canonical TATA box motifs at positions -25 to -35, relative to each of the different transcription start sites. While *RPL23aA* has only one TATA box, *RPL23aB* has TATA boxes upstream of both sets of transcription start sites (Figure 4.13). According to Perry's (2005) criteria, the TATA boxes in the *RPL23a* genes are predicted to be high-affinity binding sites for TBP (TATA-binding protein), with five or more 'preferred' nucleotides in specific positions of the seven to eight bp motif (*RPL23aA* motif, TATAaAA; *RPL23aB* motifs, TATATATt, TAaATAAA; uppercase letters, preferred nt; lowercase letters, 'acceptable' nt).

Although lacking the 5' terminal oligopyrimidine tract characteristic of vertebrate r-protein mRNAs, *RPL23aA* and *B* transcripts do have upstream leader introns (29 and 32 bp upstream of the ATG start codons, respectively) as are commonly found in the r-protein genes of mammals (Perry, 2005) and other vertebrates (Amaldi et al., 1995). Interestingly, while the 107 nt *RPL23aA* upstream intron is bounded by canonical $^{5'}$GT-AG$^{3'}$ splice sites, *RPL23aB* transcripts show splicing at non-canonical $^{5'}$TC-TA$^{3'}$ and $^{5'}$GT-TC$^{3'}$ sites, removing 214 and 215 nt introns, respectively. Given that only one 5' RACE-amplified fragment showed splicing at the $^{5'}$GT-TC$^{3'}$ sites, and that the 5' and 3' intron splice sites are flanked by direct repeats, it is probable that the $^{5'}$GT-TC$^{3'}$ processing is actually a product of missplicing. RT-PCR amplification from various wild type tissues indicated that the upstream leader introns are always processed, regardless of tissue type. What role the leader introns may play is as yet unknown, although transgenic plants carrying partial introns for each *RPL23a* gene have provided an indication of the importance of the introns for gene expression (see below).

In an effort to understand the GUS activity data for deletion series plants in light of the complex *RPL23aA* and *B* 5' UTR arrangements, RT-PCR amplifications using total RNA from deletion series plants were carried out. Amplification of 5' RR fragment::*GUS* transcripts was carried out using template from *RPL23aA* and *B*5'FR, Δ4, Δ5, and Δ6 transgenics, resulting in some surprising findings. Although *RPL23aA*Δ6 plants did not show GUS activity in any tissue examined, the transgene was transcribed in the T_2 plants, but not translated. *RPL23aA*Δ5 plants, which showed greatly reduced GUS activity compared to other *RPL23aA* deletion series T_2s, also showed transcription of the Δ5::*GUS* construct, indicating that poor translational efficiency of the transcript was responsible for weak GUS expression in these plants. The *RPL23aA*Δ5 transgene included the full 5' leader intron, but only 9 bp of sequence was left upstream of the intron 5' splice site, resulting in the observed missplicing of the intron from the Δ5::*GUS* transcript. *RPL23aAΔ6*::*GUS* transcripts lacking the 5' splice site for the leader intron, but including the 3' portion of the leader intron, were unable to be properly spliced, and, as such, the Δ6 fragment was untranslatable. In contrast, the *RPL23aA*5'FR and Δ4 plants, which showed no impairment of GUS activity, produced transgene transcripts from which the leader intron had been correctly spliced.

The results from the *RPL23aA* deletion series RT-PCRs suggest that the presence of an incorrectly spliced or incomplete leader intron led to a decrease or complete repression of translation of the GUS transcript. While the lack of a TATA box had not prevented transcription of the *RPL23aA*Δ5 and Δ6 transgenes, the lack of a correctly spliced upstream intron had clearly interfered with translation of the reporter transcript. The results of the *RPL23aB* deletion series RT-PCR, however, were rather surprising. As expected, the *RPL23aB5*'FR and Δ4 transgenes, containing the full 5' leader intron, produced properly spliced transcripts. RT-PCR from *RPL23aB*Δ5 and Δ6 plants amplified transcripts that, like *RPL23aA*Δ6 plants, were unspliced, with partial introns. Again, transcription of Δ5 and Δ6 constructs was unimpeded by lack of TATA boxes. Even though they carried unspliceable partial introns, the *RPL23aB*Δ5 and Δ6 transgenes were still able to drive *GUS* expression, indicating that, unlike *RPL23aA*Δ6, the *RPL23aB*Δ5 and Δ6 transcripts were translatable. Unexpectedly, however, ~half of

the transcripts amplified from *RPL23aB*Δ5 and Δ6 plants contained an unspliced *CATALASE* intron between the *GUS* first and second exons, the presence of which has previously been shown to prevent expression of a functional *GUS* product (Ohta et al., 1990).

Without further experimentation, it is unclear why transcription from *RPL23aA* and *B* 5' RRs could proceed without the TATA boxes found upstream of their wild type transcription start sites, and it is even less clear why the effects of an unspliceable, partial 5' leader intron were so different for the two *RPL23a* genes. Introns and intron splicing can affect gene expression in a number of different ways, at the transcriptional, post-transcriptional, or translational levels (reviewed in Le Hir et al., 2003); many of these intron-mediated mechanisms are illustrated by r-protein gene examples. As discussed above (Chapter 1, section 1.5, *ribosomal protein gene expression and regulation*), introns may harbor *cis*-regulatory elements that influence transcription, such as the YY1 element found in mammalian *L7* (Meyuhas and Klein, 1990) and *L32* (Chung and Perry, 1993). Recognition and occupation of intron splice sites by snRNAs and spliceosomal snRNPs can also increase RNA Pol II recruitment and processivity (reviewed in Le Hir et al., 2003). Introns can also play critical post-transcriptional regulatory roles via alternative splicing (e.g. mouse *S24*; Xu et al., 1994) or by the targeting of unspliced or misspliced mRNAs for degradation (e.g. Xenopus L1, Amaldi et al., 1989; mouse L32, Chung and Perry, 1989; *C. elegans L3*, *L7a*, *L10a*, and *L12*, Mitrovich and Anderson, 2000). In the case of the *RPL23a* genes, an effect of the 5' leader introns on gene regulation via transcription is unlikely, since unspliced partial introns (*RPL23aA*Δ6, *RPL23aB*Δ5 and Δ6) did not interfere with transcription of the reporter constructs. In addition, while the *RPL23aA* and *B* 5' leader introns may contain *cis* regulatory elements, any relevant predicted motifs are conspicuously absent from the introns (Figure 4.13) and deletion of portions of each intron did not eliminate or significantly repress transcription. The two most probable causes of the reduction or abolition of GUS expression in *RPL23aA* transformants carrying misspliced (Δ5) or partial (Δ6) 5' introns are increased transcript degradation or decreased translational efficiency. Intron splicing has been shown to increase the translational efficiency of numerous transcripts in animals (Matsumoto et al., 1998; Nott et al., 2004) and plants

(Bourdon et al., 2001; Rose, 2004), an effect attributed to the mRNP protein complement at exon-exon junctions (the exon junction complex, EJC) post-splicing (Nott et al., 2004). A loss of the exon-exon junction, caused by a truncated leader intron, would result in a loss of the EJC, and perhaps decrease polysome association with the mRNPs.

If the lack of correct intron splicing is the primary determinant of altered expression of the *RPL23aA*Δ5 and Δ6::*GUS* transgenes, one might expect that the partial introns of the *RPL23aB*Δ5 and Δ6 constructs would have the same effect on GUS expression. However, the *RPL23aB*Δ5 and Δ6 transgenics show almost the same expression patterns as *RPL23aB*Δ4 plants, which produce properly spliced transcripts, and no overall abolition of GUS expression is noted in any *RPL23aB* plants. One possibility for the difference between the two *RPL23a* genes is that while the *RPL23aA* upstream intron has canonical splice sites, *RPL23aB* does not, perhaps affecting mRNP protein composition pre- and post-splicing. A different mRNP composition around splice sites might also explain why ~half of the *RPL23aB*Δ5 and Δ6 transgene transcripts have an unspliced *CATALASE* intron downstream of the intact leader intron 3' splice site. Differences in intron sequence between *RPL23aA* and *B*, which share only 52% identity between their upstream leader introns, may also influence regulation. Further experimentation examining the leader introns of both genes via 3' deletions of the *RPL23aA* and *B* 5' flanking regions and intronless 5'FR constructs is currently underway and will yield more answers as to the importance of intron-mediated regulation.

CHAPTER 5. GENERAL DISCUSSION

In this thesis (Chapter 2), I have shown that an Arabidopsis r-protein L23a, AtRPL23aA, is among those r-proteins that show not only structural, but functional homology between species, via a complementation experiment. This made AtRPL23a the first plant r-protein to be confirmed as a functional member of the L23/L25 family, and the equivalent of a demonstrated multifunctional yeast protein (van Beekvelt et al., 2001; Beckmann et al., 2001; Morgan et al., 2002; Pool et al., 2002). This study, together with other heterologous complementation experiments (Rhoads and Roufa, 1987; Maki et al., 1990; Jeeninga et al., 1996; Dick and Trumpower, 1998), demonstrates the functional conservation of individual r-proteins.

Although r-proteins are highly conserved, the genes that encode them can be regulated very differently, depending on species (e.g. Mager, 1988). Prokaryotic r-protein operons are primarily controlled via autogenous regulation at the translational level (reviewed in Lindahl and Zengel, 1986; Nomura, 1999). In eukaryotes, common features of r-protein genes and transcripts contribute to their regulation. In yeast, r-protein genes are primarily regulated at the transcriptional level (Planta et al., 1995; Warner, 1999; Lieb et al., 2001), and in vertebrates, r-protein gene expression is primarily controlled at the translational level (Aloni et al., 1992; Loreni and Amaldi, 1992; Meyuhas, 2000). Unlike other eukaryotes, plants are unique in having multiple (two or more) expressed genes encoding each r-protein, even in non-polyploids like Arabidopsis (Barakat et al., 2001), often showing divergent expression patterns between genes in the same family (Williams and Sussex, 1995; Dresselhaus et al., 1999; Hughes and Friedman, 2005).

I have shown that, like other r-protein genes, the genes encoding Arabidopsis RPL23a are expressed in mitotically active tissues; however, *RPL23aA* and *B* are also differentially expressed (Chapter 3). *RPL23aB* transcript levels were generally lower than those of *RPL23aA*, and transcript levels for the two genes responded differently to stresses. The expression patterns for the *RPL23a* genes under various stress conditions

also differed from those reported for other r-protein genes (Sáez-Vásquez et al., 2000; Ludwig and Tenhaken, 2001; Kim et al., 2004). While the 5' regulatory regions for the two *RPL23a* genes showed numerous putative regulatory motifs, it was unclear whether these predicted elements are functional.

To evaluate the link between putative motifs and functional *cis* regulatory elements controlling *RPL23aA* and *B* expression, a 5' regulatory region analysis was carried out via a 5' deletion series and transcript analysis for each gene (Chapter 4). Plants carrying the full 5' RR for each gene confirmed and expanded the differential expression profiles for *RPL23aA* and *B* in an array of untreated tissues, but serial deletions of the 5' RRs did not clearly correlate with tissue-specific expression in almost all cases. The final two deletions of the *RPL23aA* 5'RR (145 to 66 bp of 5'RR) resulted in a severe reduction in activity, and removal of a 102 bp region between the third and fourth *RPL23aB* deletion fragments resulted in a loss of activity in anthers and pollen. As with a previous deletion series study in tobacco (*L34*; Dai et al., 1996), the observed deletion series expression patterns did not correlate well with putative motifs identified in the 5'RRs of *RPL23aA* and *B*. While *telo* boxes (Lenvik et al., 1994; Trémousaygue et al., 1999) and two *PCNA* site II motifs (Trémousaygue et al., 2003; upstream of the *telo* boxes in *RPL23aB*), common elements identified in plant r-protein genes, were identified in the *RPL23a* genes, it is unclear what role these elements may have in expression of the two genes. Like the case in tobacco *L34* (Dai et al., 1996), no predicted wound-responsive elements were identified in the *RPL23a* genes, although the genes clearly contain wound-inducible elements, and while both genes were strongly expressed in root, elimination of the predicted root-specific motifs did not abolish this activity. In addition, elimination of the same motifs from both RPL23a genes did not have the same effect on expression; for example, elimination of all but one putative pollen-specific motif from the *RPL23aA* 5'RR did not reduce expression in pollen (Δ4 transgenics), whereas the same loss in *RPL23aB* plants (Δ4) abolished pollen-specific expression.

The lack of a correlation between the 5' deletions and predicted *cis* motifs is probably due to the fact that transcription of *RPL23aA* and *B* did not appear to be the primary regulatory factor controlling gene expression. Not only did the expression studies using wild type tissues (Chapter 3) show quite subtle differences in transcript

levels between tissues and genes, RT-PCR from 5' deletion series transgenics confirmed the presence of transcripts for each transgene, indicating that the weak or abolished expression seen in *RPL23aA*Δ5 and Δ6 transgenics was due to a repression of translation (or post-transcriptional processes such as mRNA transport out of the nucleus), not transcription. Missplicing or lack of splicing of the 5' leader introns in *RPL23aA* transgene mRNAs was correlated with reduced GUS expression, and although *RPL23aB* transgene transcripts containing unspliced, partial 5' leader introns did not appear to have impaired expression of the reporter gene, ~half the transcripts contained unspliced *CATALASE* introns within the *GUS* gene.

As more data are compiled regarding r-protein gene structure and expression, it is becoming increasingly obvious that introns are a widespread feature of eukaryotic r-protein gene composition. R-protein genes account for a third of the intron-containing genes in the entire yeast genome (Planta and Mager, 1998; Spingola et al., 1999) and ~two thirds of the intron-containing genes in *Guillardia theta* (Grossman, 2005). The intron-containing r-protein genes of yeast (Planta and Mager, 1998), *G. theta* (Grossman, 2005), Xenopus (Amaldi et al., 1995), mammals (Perry, 2005), and plants (Giannino et al., 2000) all tend to have introns near or in the 5' regulatory region, a position well documented to positively influence gene expression (Matsumoto et al., 1998; Bourdon et al., 2001; Rose, 2004). Although 5' *cis* regulatory motifs (e.g. Rap1 in yeast, YY1 in mammals, *telo* boxes in plants) may differ between species, the importance of introns in r-protein gene composition appears to have remained constant. While the presence of introns may simply reflect the evolutionary age of r-protein genes (Poole et al., 1999), it is also possible that introns are a key regulatory feature. In order to elucidate the importance of the 5' leader introns in *AtRPL23a* gene expression, I have designed a series of 3' deletions and intronless fragments of the *RPL23aA* and *B* 5' RRs, which have been used to generate transgenics currently being studied by others in the Bonham-Smith lab. The key to understanding the role 5' leader introns may play in coordinate plant r-protein gene regulation, however, requires a survey of all identified plant r-protein genes. This survey would be most easily carried out using the list of r-protein genes for Arabidopsis (Barakat et al., 2001), most of which have been experimentally confirmed as ribosomal constituents (Chang et al., 2005). A comparison

of cDNA and genomic sequences for each Arabidopsis r-protein gene, such as that reported by Perry (2005) for mammals, would yield information about mRNA processing for each transcript.

The differences between the *RPL23aA* and *B* 5'RRs, in terms of both primary sequence and functional elements as demonstrated by the expression and deletion series studies, are surprising given the requirement for coordinate expression of r-protein genes, and indicate that sequence comparison alone can be misleading. In other eukaryotes such as yeast or mammals, shared characteristics between r-protein gene regulatory regions, either in terms of *cis* elements or 5' UTR structure, appear to be the key to coordinate expression. Despite the divergence between *RPL23a* 5' RRs, however, the expression patterns for the two genes have remained remarkably similar with few major differences. Given the importance of post-transcriptional and/or translational processes in the regulation of the *RPL23a* genes, a future study of *RPL23aA* and *B* transcripts in terms of hnRNP composition, localization, and polysome association would help elucidate control mechanisms for each gene.

While it is unlikely that the differences in *RPL23aA* and *B* gene expression reflect extraribosomal functions for one of the isoforms, since RPL23a is among the most highly conserved core r-proteins, it is possible that differences in expression for the two genes reflects heterogeneity in ribosome composition as observed for other Arabidopsis r-proteins (Chang et al., 2005). The function of the plant-specific phenomenon of multiple expressed isoforms for each r-protein, the need for ribosomal heterogeneity, the coordination of expression between r-protein genes, and even the primary mechanism by which plant r-protein genes are regulated, are all open questions that are only beginning to be answered.

APPENDIX 1: PEPTIDE BOND FORMATION AND PEPTIDYL TRANSFERASE

Translation is the process by which polypeptides are produced according to an mRNA template; this process proceeds through initiation, elongation, and termination phases. Elongation is the phase of translation where the ribosome catalyzes peptide bond formation in cycles that involve three major steps: binding of an incoming aminoacyl tRNA with an anticodon matching the mRNA codon exposed in the A site of the small subunit, peptide bond formation between amino acids linked to tRNAs in the A and P sites, and translocation of the A-site tRNA to the P site (moving the mRNA one codon over) while the deacylated P site tRNA relocates to the E site. The second elongation step, peptide transfer, is catalyzed by the ribosome in the absence of external protein factors. This appendix details the current knowledge of the peptidyl transferase function of the ribosome.

Amino acids are bound to tRNAs via ester linkages formed between the α-carboxyl group of the amino acid and the 3' hydroxyl group of the adenine (A76) of the CCA sequence found at the 3' terminus (on the acceptor stem) of all tRNA molecules. Peptide bonds are formed within the large subunit of the ribosome between an amino acid linked to a tRNA in the A site (the aminoacyl tRNA, the A site substrate) and an amino acid/peptide linked to a tRNA in the P site (the peptidyl tRNA, the P site substrate). Peptidyl transferase is the catalysis of peptidyl transfer - the nucleophilic attack of the α-amino group of the A site substrate on the ester carbonyl group (the carbonyl carbon of the ester linkage between an amino acid/peptide and the peptidyl tRNA) of the P-site substrate (Figure A1.1); the ribosome increases the rate of this reaction by at least 10^5- (Rodnina and Wintermeyer, 2003) to 10^7-fold (Sievers et al., 2004). How peptidyl transfer is catalyzed by the ribosome is the focus of much current research; structural, biochemical, and reaction kinetics studies are being used in prokaryotic systems to investigate the active site mechanics. Recent studies suggest that peptidyl transfer catalysis results from a combination of substrate positioning (Nissen et

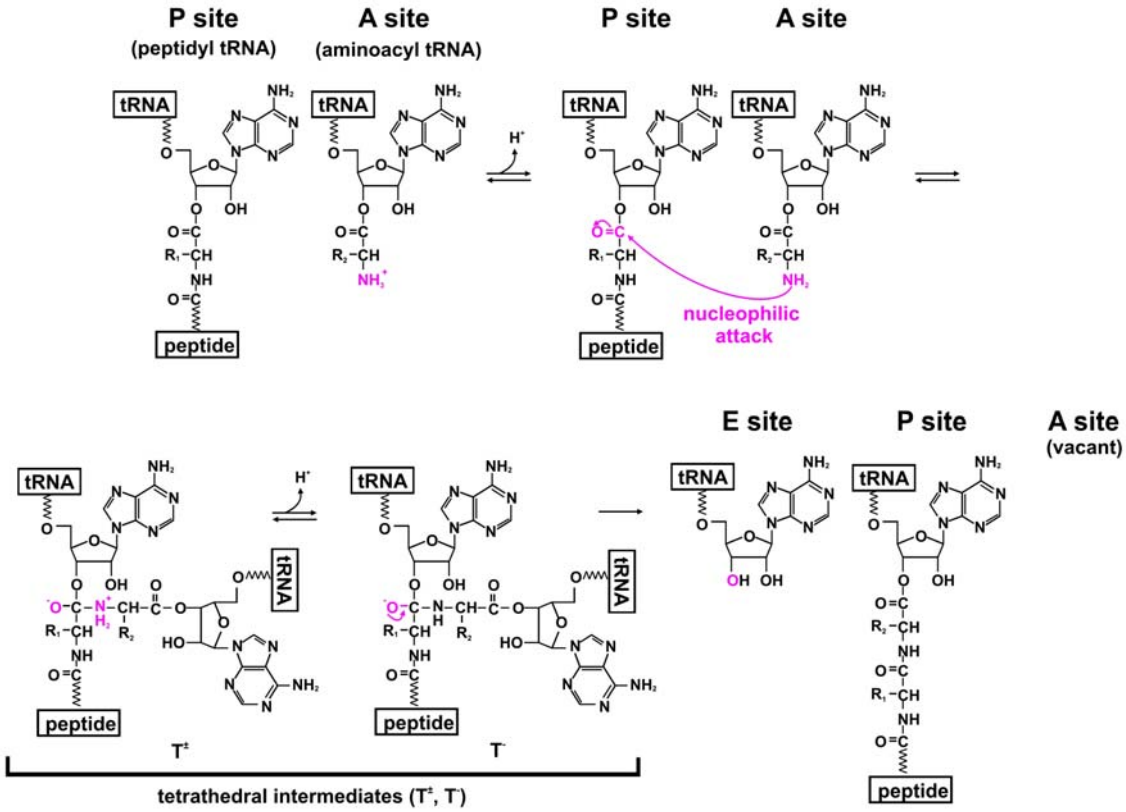


Figure A1.1. Peptide bond formation. First step, deprotonation of the α-amino group (NH_3^+, **pink**) of the amino acid carried by the aminoacyl tRNA (A site), followed by the second step, nucleophilic attack by the deprotonated α-amino group (NH_2) on the carbonyl carbon of the the ester linkage between the nascent peptide and the peptidyl tRNA (P site). This is followed by the formation of the zwitterionic tetrahedral intermediate ($T^{\pm}$), which is then deprotonated to form a negatively charged intermediate (T^-). Protonation of the leaving oxygen (**pink** in last step) results in product formation – deacylated P site tRNA, which moves into the E site, and new peptidyl tRNA which moves from the A site to the P site, leaving the A site vacant. R_n, amino acid R group (side chain). Modified from Rodnina and Wintermeyer (2003).

al., 2000; Hansen et al., 2002; Sievers et al., 2004) and chemical catalysis (Katunin et al., 2002; Moore and Steitz, 2002; Steitz and Moore, 2003; Weinger et al., 2004; Steitz, 2005).

Structural studies of prokaryotic ribosomes have demonstrated that peptidyl transferase and substrate positioning are functions exclusively of RNA (Ban et al., 2000; Nissen et al., 2000). The central loop of domain V of the 23S rRNA is the peptidyl transferase center (PTC) and domain V nucleotides 3' and 5' to the central loop form the A and P sites of the large subunit; the closest r-protein side-chain extensions to the PTC are ~18 Å away (Nissen et al., 2000). Crystal structures of the *H. marismortui* 50S subunit complexed with substrate analogues, including analogues of the tetrahedral intermediate formed between the aminoacyl carbonyl carbon and the peptidyl tRNA, indicate tRNA terminal CCA sequence interactions with the LSU rRNA at their respective A and P sites (Nissen et al., 2000; Hansen et al., 2002). In the P site, analog residues corresponding to peptidyl tRNA C74 and C75 base pair with G2285 and G2284 (G2252 and G2251 in *E. coli*) of the LSU rRNA P loop (Nissen et al., 2000; Hansen et al., 2002). In the A site, C75 of the aminoacyl tRNA analogue forms a base pair with G2588 (G2553) of the A-loop (Nissen et al., 2000; Hansen et al., 2002). The interaction of tRNA substrates with the rRNA of the LSU A and P sites positions the substrates for activity (Nissen et al., 2000; Hansen et al., 2002). From measurements of activation parameters for uncatalyzed and ribosome-catalyzed peptide transfer (Sievers et al., 2004), it has been suggested that catalysis of peptidyl transfer by the ribosome is achieved solely by correct binding and alignment of substrates; however, other evidence suggests that substrate alignment is not the sole criteria by which the ribosome increases reaction rate (see below).

Whether ribosomal peptidyl transferase involves chemical catalysis has been a difficult question to answer. It is known that peptidyl transferase activity is at least partly pH-dependent (Maden and Monro, 1968). A ribosomal group with a pK_a (acid dissociation constant) of 7.5 is involved in the pH-dependent reaction; protonation of this single ionizing group resulted in a 100-fold reduction in reaction rate (Katunin et al., 2002). Initially it was thought that N3 of A2486 (A2451 in *E. coli*) of the *H. marismortui* 50S active site might be such an ionizing group, acting as a general base to

catalyze the nucleophilic attack of peptide transfer, but this would require a pK_a greater than 7, much higher than the usual value for N3 of adenine (pK_a ~1.5; Nissen et al., 2000). It was proposed that this pK_a increase could be achieved by a hydrogen bonding 'charge relay' network between A2486, G2482 (G2447 in *E. coli*), and a buried phosphate of A2485 (A2450), increasing the negative charge on N3 of A2486 (Nissen et al., 2000; Hansen et al., 2002; Moore and Steitz, 2002; Steitz and Moore, 2003). Substitution of A2451 (G, U, or C) and G2447 (A, U, or C) in *E. coli* 23S rRNAs, followed by functional studies, resulted in decreased peptidyl transferase activity (Polacek et al., 2001). Similarly, an A2451U mutation inhibited peptidyl transferase activity by ~130-fold and abolished pH sensitivity due to protonation of a ribosomal group (Katunin et al., 2002). It is as yet unknown what the role of A2451 might be; even if A2451 is the pH-dependent constituent of the peptidyl transferase, it is unclear whether pH dependence is due to acid/base catalysis, or if there is a pH-dependent conformational change of the active site that affects A2451, as well as other bases (Bayfield et al., 2001**;** Rodnina and Wintermeyer, 2003). Certainly, a number of factors are involved, including positioning of substrates at the active site (see above); even protonated ribosomes show reaction rates $>10^3$ times faster than uncatalyzed peptide transfer (Katunin et al., 2002).

In addition to substrate positioning and pH-dependent functions at the ribosomal active site, substrate-mediated catalysis also represents a major contribution to peptidyl transferase activity (Weinger et al., 2004). An *in vitro* puromycin-saturated system (post-translocation complexes of *E. coli* 70S ribosomes with radiolabeled fMet-Phe-$tRNA^{Phe}$ in the P site + excess puromycin→fMet-Phe-puromycin) was used where the limiting step of peptide bond formation was catalysis; this reaction proceeds at a maximum rate of ~50 bonds·s^{-1} at physiological pH, with native tRNA occupying the P-site (pH ≥ 7.7; Katunin et al., 2002; Weinger et al., 2004). Modeling of the 50S subunit complexed with A- and P-site substrate analogs indicated that the 2' hydroxyl of peptidyl tRNA A76 forms a hydrogen bond with the attacking aminoacyl α-amino group, positioning A76 to aid catalysis (Hansen et al., 2002). The importance of the P-site A76 2'OH was confirmed by substitution of A76 with 2'-deoxy-A (dA76) and 2'-deoxy-2'-fluoro-A (fA76); removal of this 2'OH resulted in a 10^6-fold reduction in

reaction rate compared to A76 tRNA (Weinger et al., 2004). Clearly the A76 2'OH plays a critical role in peptidyl transfer; it may act as a general acid/base to remove a proton from the attacking α-amino group and donate a proton to the leaving (deacylated) 3'O group (Weinger et al., 2004; Steitz, 2005). The structural change associated with dA76/fA76 may have been responsible for the reduced rate of catalysis (Steitz, 2005); however, it seems unlikely given the equivalent effects of both substrates. More biochemical and structural data is required before the mechanism of peptidyl transferase is fully understood.

APPENDIX 2: RIBOSOMAL PROTEIN HOMOLOGUES AND FUNCTIONS

The following table is a compilation of information about r-protein families across all domains of life, with *E. coli*, Arabidopsis, rat, and yeast homologues listed for each family. This table is a guide to r-protein homologues across species with any known functional data associated with each entry in order to facilitate connections between homologues and r-proteins from related subunit regions. More r-protein (extraribosomal) functions, primarily for prokaryotic r-proteins, are listed in Wool (1996). The following legend (pp. 170-171) applies to the table in full.

Table A2.1. R-protein homologues, rRNA contacts, and known functions. R-protein gene families grouped according to conservation across all three domains of life (Bacteria, Archaea, Eukarya), across two domains, or unique to single domains; r-protein family name and phylogenetic grouping given as in Lecompte et al. (2002) unless otherwise specified. Capital letters following a eukaryotic gene name indicate multiple expressed isoforms. *, S15a copies A,C,D,F according to Hulm et al., 2005; S15aB and E are mitochondrial. ** recently confirmed as an r-protein (Sengupta et al., 2004). *** plant-specific r-protein (Szick et al., 1998; Barakat et al., 2001). –, indicates no homologue for a given family, or no known rRNA domain interaction; blank spaces indicate unknown rRNA contacts/functions.

rRNA secondary structure domain contacts given as for bacterial (Brodersen et al., 2002) or archaebacterial (Klein et al., 2004) family members, as determined by 30S and 50S crystal structures. SSU rRNA: 3'Ma, 3' major domain; 3' mi, 3' minor domain; 5', 5' domain; B, part of the body structural feature of the 30S subunit; C, central domain; D, part of the decoding region, between the head and body of the 30S subunit; H, part of the head of the 30S subunit; P, part of the platform of the 30S subunit. LSU: contacts listed with domains I-VI or 5S rRNA.

Functions are listed as determined by structural and functional studies in a variety of species. A, assembly, folding/initiation; anti., transcription antitermination;

Ca, Ca^{2+}-dependent calmodulin binding protein; cellcyc., cell cycle/apoptosis; Cp, ceruloplasmin regulation/defence response; D, decoding center contact with A- and/or P-site; DB anemia, Diamond-Blackfan anemia; dev, development; ET, peptide exit tunnel; FB, factor binding (sarcin-ricin domain, SRD); gen. stress, genotoxic stress response; ISB, inter-subunit bridge; *Minute* – *Minute* phenotype; S, structural role, rRNA stabilization; S/5S-23S, involved in attachment of 5S rRNA to 23S rRNA; S/E site, structural role as part of E site, factor-binding center; S/LP, structural role as part of lateral protuberance, L1 arm or L7/L12 stalk; selenocys, selenoprotein mRNA translational recoding; signal trans., signal transduction; SUint, intersubunit interface.

References:1, Lecompte et al., 2002, supplementary material table S2; **2**, Barakat et al., 2001; **3**, Wool et al., 1995; **4**, Planta and Mager, 1998; **5**, Brodersen, 2002; **6**, Ban, 2000; **7**, Klein, 2004; **8**, Jagannathan and Culver, 2003; **9**, Nikulin et al., 2003; **10**, Yusupov et al., 2001; **11**, Gao et al., 2003; **12**, Spahn et al., 2001; **13**, Halic et al., 2005; **14**, Berisio et al., 2003; **15**, Woolhead et al., 2004; **16**, Miyake et al., 2005; **17**, Andersson et al., 1994; **18**, Van Lijsebettens et al., 1994 ; **19**, Nishimura et al., 2004; **20**, Sonnemann et al., 1991; **21**, Shen et al., 2005; **22**, Naora and Naora, 1999; **23**, Revenkova et al., 1999; **24**, Nilsson et al., 2004; **25**, Chavatte et al., 2005; **26**, Mazumder et al., 2003; **27**, Torres et al., 2001.

Phylogenetic Distribution[1]	r-protein Family[1]	*Arabidopsis RP*[2,3] *(Rattus norvegicus* nomenclature*)*	*S. cerevisiae* homologue[1,4]	*E.coli* homologue[1]	Domain contact(s)	Function	Reference
Bacteria, Archaea, Eukarya	S2p	SaA-B	S0A-B	S2	3'Ma, C (H/B)	S	5
	S3p	S3A-C	S3	S3	3'Ma, 5' (H)	S	5
	S4p	S9A-C	S9A-B	S4	C, 5' (D)	S, A, anti	5, 27
	S5p	S2A-D	S2	S5	3'Ma, C, 5' (D)	S	5
	S7p	S5A-B	S5	S7	3'Ma, C, (H)	S, A	5
	S8p	S15aA,C,D,F*	S22A-B	S8	C, 5' (B)	S, A, Ca	5, 8, 21
	S9p	S16A-C	S16A-B	S9	3'Ma (H)	D	5
	S10p	S20A-C	S20	S10	3'Ma (H)	S	5
	S11p	S14A-C	S14A-B	S11	C, 3'mi (B, P)	S, Ca	5, 21
	S12p	S23A-B	S23A-B	S12	C, 5', 3' mi (D)	S, D	5
	S13p	S18A-C	S18A-B	S13	3'Ma (H)	SUint, D, ISB, dev	5, 10, 11, 18
	S14p	S29A-D	S29A-B	S14	3'Ma (H)	S	5
	S15p	S13A-B	S13	S15	C (B)	S, A, ISB	5, 8, 10, 11
	S17p	S11A-C	S11A-B	S17	5', C (B)	S, A	5
	S19p	S15A-F	S15	S19	3'Ma (H)	SUint, ISB	5, 11

Table A2.1. See legend, pp. 170-171.

Phylogenetic Distribution[1]	R-protein Family[1]	*Arabidopsis* RP[2,3] *(Rattus norvegicus* nomenclature*)*	*S. cerevisiae* homologue[1,4]	*E.coli* homologue[1]	Domain contact(s)	Function	Reference
Bacteria, Archaea, Eukarya	L1p	L10aA-C	L1A-B	L1	V	S/LP	6, 9
	L2p	L8A-C	L2A-B	L2	II, III, IV, V	S, A, ISB	7, 10, 11
	L3p	L3A-C	L3	L3	II, IV, V, VI	FB, S, A, anti	6, 7, 27
	L4p/L4e	L4A-D	L4A-B	L4	I, II, V	S, A, ET, anti	7, 15, 27
	L5p	L11A-D	L11A-B	L5	V, 5S	S/5S-23S, ISB	6, 7, 10, 11
	L6p	L9A-D	L9A-B	L6	II, V, VI	FB	6
	L10p	P0A-C	P0 (A0)	L10	II	S/LP, FB	6
	L11p	L12A-C	L12A-B	L11	II	S/LP, FB	6
	L12p	P1A-C, P2A-E	P1A-B, P2A-B	L12	-	S, LP, FB	6
	L13p	L13aA-D	L16A-B	L13	II, V, VI	FB, Ca, Cp, anti	6, 21, 26, 27
	L14p	L23A-C	L23A-B	L14	IV, V, VI	FB, ISB, cellcyc	6, 7, 10, 11, 22
	L15p	L27aA-C	L28	L15	I, II, V	S	7
	L18p	L5A-C	L5	L18	II, V, 5S	S/5S-23S	6, 7
	L22p	L17A-B	L17A-B	L22	I, II, III, IV, V, VI	S, A, ET	6, 7, 14, 15
	L23p	L23aA-B	L25	L23	I, III	ET	6, 7

Table A2.1. See legend, pp. 170-171.

Phylogenetic Distribution[1]	R-protein Family[1]	*Arabidopsis* RP[2,3] *(Rattus norvegicus* nomenclature*)*	*S. cerevisiae* homologue[1,4]	*E.coli* homologue[1]	Domain contact(s)	Function	Reference
Bacteria, Archaea, Eukarya	L24p	L26A-B	L26A-B	L24	I	ET, S, A	6, 7
	L29p	L35A-D	L35A-B	L29	I	ET	6, 7
	L30p	L7A-D	L7A-B	L30	II, 5S	S/5S-23S	6, 7
	L7ae	L7aA-B	L8A-B	-	I	E site (?), ISB(?)	6
Archaea, Eukarya	S3ae	S3aA-B	S1A-B	-		*Minute*, cellcyc	17, 22
	S4e	S4A-D	S4A-B	-		Ca	21
	S6e	S6A-B	S6A-B	-			
	S8e	S8A-B	S8A-B	-		Ca	21
	S17e	S17A-D	S17A-B	-			
	S19e	S19A-C	S19A-B	-		DB anemia	16
	S24e	S24A-B	S24A-B	-			
	S25e	S25A-E	S25A-B	-			
	S26e	S26A-C	S26A-B	-			
	S27ae	S27aA-C	S31	-		gen. stress	23
	S27e	S27A-C	S27A-B	-			
	S28e	S28A-C	S28A-B	-			
	S30e	S30A-C	S30A-B	-			

Table A2.1. See legend, pp. 170-171.

Phylogenetic Distribution[1]	R-protein Family[1]	*Arabidopsis* RP[2,3] *(Rattus norvegicus* nomenclature*)*	*S. cerevisiae* homologue[1,4]	*E.coli* homologue[1]	Domain contact(s)	Function	Reference
Archaea, Eukarya	L10e	L10A-C	L10	-	II, V, 5S	S/5S-23S	6, 7
	L13e	L13A-D	L13A-B	-			
	L14e	L14A-B	L14A-B	-			
	L15e	L15A-B	L15A-B	-	I, II, III, IV, V	S	7
	L18e	L18A-C	L18A-B	-	II	S	6, 7
	L19e	L19A-C	L19A-B	-	II, III, IV, VI	ET, Ca	6, 7, 20
	L21e	L21A-F	L21A-B	-	II, V, 5S	S/5S-23S	6, 7
	L24e	L24A-B	L24A-B	-	IV, VI	FB, ISB, dev	6, 7, 12, 19
	L30e	L30A-C	L30	-		ISB, selenocys	13, 25
	L31e	L31A-C	L31A-B	-	III, IV, VI	ET	6, 7
	L32e	L32A-B	L32	-	I, II	S	6, 7
	L34e	L34A-C	L34A-B	-			
	L35ae	L35aA-D	L33A-B	-			
	L37ae	L37aA-C	L43A-B	-	II, III, IV	ISB	7, 12
	L37e	L37A-C	L37A-B	-	I, II, III, IV	S, ET	6, 7

Table A2.1. See legend, pp. 170-171.

Phylogenetic Distribution[1]	**R-protein Family**[1]	***Arabidopsis* RP**[2,3] ***(Rattus norvegicus* nomenclature*)***	***S. cerevisiae* homologue**[1,4]	***E.coli* homologue**[1]	**Domain contact(s)**	**Function**	**Reference**
Archaea, Eukarya	L38e	L38A-B	L38	-			
	L39e	L39A-C	L39	-	I, III	ET	6, 7, 15
	L40e	L40A-B	L40A-B	-			
	L41e	L41A-G	L41A-B	-			
	L44e	L36aA-B	L42A-B	-	I, IV, V	S/E site	6, 7
Eukarya	S7e	S7A-C	S7A-B	-			
	S10e	S10A-C	S10A-B	-			
	S12e	S12A-C	S12	-			
	S21e	S21A-C	S21A-B	-			
	RACK1**	RACK1	Asc1p/RACK1	-		signal trans.	24
	L6e	L6A-C	L6A-B	-			
	L18ae	L18aA-C	L20A-B	-			
	L22e	L22A-C	L22A-B	-		Ca	21
	L27e	L27A-C	L27A-B	-			
	L28e	L28A-C	-	-			
	L29e	L29A-B	L29	-			
	L36e	L36A-C	L36A-B	-			
		P3A-B***	-	-			

Table A2.1. See legend, pp. 170-171.

Phylogenetic Distribution[1]	R-protein Family[1]	*Arabidopsis* RP[2,3] *(Rattus norvegicus* nomenclature*)*	*S. cerevisiae* homologue[1,4]	*E.coli* homologue[1]	Domain contact(s)	Function	Reference
Bacteria	S1p	-	-	S1			
	S6p	-	-	S6	C (C, P)	S	5
	S16p	-	-	S16	5', C (5'/B)	S	5
	S18p	-	-	S18	C (C, P)	S	5
	S20p	-	-	S20	5'/3'mi (5'/B)	S	5
	S21p	-	-	S21			
	S22p	-	-	S22			
	S31e	-	-	-			
	L9p	-	-	L9			
	L16p	-	-	L16			
	L17p	-	-	L17			
	L19p	-	-	L19		ISB	10, 11
	L20p	-	-	L20			
	L21p	-	-	L21			
	L25p	-	-	L25			
	L27p	-	-	L27			
	L28p	-	-	L28			
	L31p	-	-	L31			

Table A2.1. See legend, pp. 170-171.

Phylogenetic Distribution[1]	**R-protein Family**[1]	***Arabidopsis* RP**[2,3] **(*Rattus norvegicus* nomenclature)**	***S. cerevisiae* homologue**[1,4]	***E.coli* homologue**[1]	**Domain contact(s)**	**Function**	**Reference**
Bacteria	L32p	-	-	L32			
	L33p	-	-	L33			
	L34p	-	-	L34			
	L35p	-	-	L35			
	L36p	-	-	L36			

Table A2.1. See legend, pp. 170-171.

APPENDIX 3. A RAPID *AGROBACTERIUM*-MEDIATED *ARABIDOPSIS THALIANA* TRANSIENT ASSAY SYSTEM

Stable transformation of *Arabidopsis thaliana* is a lengthy process involving up to three months of plant growth and seed selection. We have developed a rapid (three-week) transient assay system to test the functionality of *cis*-regulatory regions controlling expression of a reporter gene in plants before undertaking stable transformation. Two-week old Arabidopsis seedlings were vacuum-infiltrated with *Agrobacterium tumefaciens* cultures carrying various upstream regulatory regions controlling *uidA* (*β-glucuronidase*, *GUS*) expression. Seedlings were fixed and stained for GUS activity three to five days following infiltration. Regulatory regions tested in this system include the *CaMV 35S* promoter, the upstream regulatory region of ribosomal protein gene *L23aA*, and a temperature-inducible regulatory region (*HSP101B*) also from Arabidopsis. The percentage of seedlings positive for GUS activity varied depending on the construct used, with the *CaMV 35S* promoter producing the highest numbers of GUS-positives. Temperature induction treatments elicited increased GUS expression in seedlings transformed with the *HSP101B* regulatory region. Regardless of construct, GUS expression levels were higher in seedlings collected five days following *Agrobacterium* infiltration than those collected at three or four days post-infiltration.

Introduction

Transient assay systems have proven useful in a wide array of research by providing an alternative to stable transformation. Popular transient assay techniques include the use of protoplasts transformed via electroporation or polyethylene glycol-mediated techniques (Sheen, 2001), and particle bombardment of epidermal cells or leaves (e.g. Sessa et al., 1998; Schweizer et al., 1999). While both protoplast and biolistics techniques have proven useful in some species, protoplast cultures can be difficult and time-consuming to maintain, and particle bombardment requires specialized

equipment. It is important to develop a rapid, economical transient assay system focused on Arabidopsis, a key model plant, as heterologous systems may not reflect the true activity of Arabidopsis genes.

Agrobacterium tumefaciens is a readily available tool for the production of transgenic plants. The few transient assay systems that focus on *Agrobacterium* infiltration as a transformation method include tobacco seedling (Rossi et al., 1993), tobacco leaf (Yang et al., 2000), and mature Arabidopsis (Rakouský et al., 1997) assays. The *Agrobacterium*-mediated transient assay system presented here uses Arabidopsis seedlings, bypassing the time required to generate mature plants. Most previously developed *Agrobacterium*-mediated transient assay systems have used strong constitutive promoter constructs, such as *CaMV 35S::GUS* (Rossi et al., 1993; Rakouský et al., 1997). The system presented here was tested using multiple constructs incorporating strong, weak, and inducible regulatory regions to drive reporter gene expression. Testing constructs carrying a variety of regulatory region types is an important step in evaluating the overall effectiveness of the transient assay system.

Materials and Methods

Plant material and seedling cultivation

Arabidopsis thaliana (ecotype Columbia) seedlings were used in the transient assay experiments. Seed was sterilized overnight (18-20 hours) using a vapor-phase sterilization method (Clough and Bent, 1998). Seedlings were grown on ½ Murashige and Skoog medium (MS; Murashige and Skoog, 1962; Sigma, St. Louis, MO) containing 15 gL^{-1} sucrose and $8gL^{-1}$ Phytagar (Invitrogen, Carlsbad, CA) in petri plates for two weeks at 23°/18°C, 16/8 h day/night cycle.

Strains and plasmids

Four different constructs were used in the study. The AtRPL23aA construct was assembled as follows: a 1.5 kb region immediately upstream of the Arabidopsis ribosomal protein (RP) *L23aA* open reading frame was amplified from a BAC (F12L6, GenBank accession no. AC004218) and cloned into the binary vector pCAMBIA1381Z (CAMBIA, Canberra, Australia) upstream of the *uidA* (*GUS*) gene. A *CATALASE* intron is present in the *GUS* gene in the CAMBIA vectors, thereby preventing GUS expression in bacteria (Ohta et al., 1990). The HSP101B construct contains a 2 kb fragment of the 5'

regulatory region of a *HSP101* pseudogene (within a *MULE-24* transposon; Yu et al., 2000) amplified from an Arabidopsis BAC (F3D13, GenBank accession no. AF069300) and cloned 5' to the *GUS* gene in pCAMBIA1381Z. The fragment, *HSP101B*, contains one consensus heat shock element (HSE; alternating repeats of $^{5'}$nGAAn$^{3'}$; Barros et al., 1992; Schöffl et al., 1998), one HSE with a single nucleotide change (compared to *HSP101*, accession no. AF329939; Hong and Vierling, 2001), and four conserved Cor15a-like low-temperature response elements (LTREs; $^{5'}$CCGAC$^{3'}$; Baker et al., 1994; Jiang et al., 1996). pALPHONOS consists of the 2.3 kb regulatory sequence immediately upstream of the Arabidopsis *LEAFY* ORF (accession no. M91208; Weigel et al., 1992) cloned 5' to the Arabidopsis *HSP101* ORF (accession no. U13949; Schirmer et al., 1994), cloned into pCGN1558 (McBride and Summerfelt, 1990). No *GUS* gene is present in the construct, designating pALPHONOS as a negative control. pCAMBIA2301, containing the *GUS* gene under the control of the *CaMV 35S* promoter, was used as a positive control. *Agrobacterium tumefaciens* strain LBA4404 carrying the pAL4404 vir-containing plasmid (Hoekema et al., 1983) was used as the host for all constructs.

Agrobacterium *infiltration*

To prepare the infiltration medium, 250 mL LB broth was inoculated with *Agrobacterium* stock (1 mL 15% glycerol *Agrobacterium* stock). LB was supplemented with the appropriate antibiotic (*pALPHONOS*: gentamicin 20 μgmL^{-1}, streptomycin 100 μgmL^{-1}; all other cultures: kanamycin 25 μgmL^{-1}, streptomycin 100 μgmL^{-1}). Cultures were grown at 28°C for ~24 h with shaking (250 rpm). *Agrobacterium* cells were collected by centrifugation (5000 rpm/3836 *g* for 10 minutes at room temperature, Beckman JA-14 rotor) then resuspended in 5% sucrose, 0.01% Silwet-L77 as per Clough and Bent (1998) to a final OD_{600} of 1.1 - 1.3. Preliminary experiments using *Agrobacterium* suspensions at an OD_{600} of 0.7 - 0.8 showed much lower transformation efficiencies than replicates using cultures with an OD_{600} of >1.

Seedlings were infiltrated by pouring 25-35 mL (enough to submerge the seedlings) of infiltration medium onto each plate and subjecting the plates to 70-100 kPa (~25 mm Hg) vacuum for two minutes. Excess infiltration medium was poured off and plates were re-sealed with plastic wrap. The plates were placed right side up in a growth

chamber for three to five days (23°/18°C, 16/8 h day/night cycle). Controls for each replicate of the transient assay included vacuum-only, sucrose/Silwet-L77 solution only, and *pALPHONOS* infiltration negative control.

Induction treatments and GUS histochemical assay

Preliminary experiments where seedlings were collected at 24 and 48 h following infiltration showed little to no GUS expression; GUS activity was therefore examined at three later time points: 72, 96, and 120 h (three to five days). One hour prior to seedling collection at each time point, two plates of *HSP101B*-infiltrated seedlings were heat treated at 35°C and two plates were cold treated at 15-17°C, in growth chambers, for 60 minutes.

Two plates of seedlings were collected from each treatment at each time point. Seedlings were removed from plates and placed in vials for staining. X-gluc staining was carried out as modified from Sieburth and Meyerowitz (1997)**.** Tissues were fixed in 90% acetone on ice for 15-20 minutes, then rinsed in a solution of 50 mM $NaPO_4$, pH 7.2, 0.5 mM $K_3Fe(CN)_6$, and 0.5 mM $K_4Fe(CN)_6$ for ~5 minutes. After rinsing, X-gluc staining solution (50 mM $NaPO_4$, pH 7.2, 2 mM X-gluc [5-bromo-4-chloro-3-indolyl β-D-glucuronide; Rose Scientific, Alberta, Canada], 0.5 mM $K_3Fe[CN]_6$, 0.5 mM $K_4Fe[CN]_6$) was added to vials and vacuum infiltration of seedlings was carried out for two minutes (70-100 kPa/~25 mm Hg). Seedlings were incubated in X-gluc staining solution for 48 h at 37°C after which chlorophyll was removed from seedlings with three changes of 70% ethanol. GUS activity was scored on the basis of X-gluc staining visualized under a stereomicroscope (Wild M3Z, Wild Heerbrugg). Percentage of GUS-positives was calculated based on how many seedlings showed X-gluc staining compared to unstained seedlings. The transient assay was tested in triplicate using all constructs and controls.

Results

Identification of GUS-positive seedlings as a function of time

GUS activity was observed in cotyledons (Figure A3.1), leaves, and/or roots of some seedlings transformed with *AtRPL23aA*, *HSP101B*, and *pCAMBIA2301* constructs. An average of 1502 seedlings was counted per construct per time point, with the mean percentage of seedlings showing GUS activity ranging from ~0.3 to 30.5%. No GUS

activity was detected in any of the controls at any time point over all three replicates. For all constructs tested an increase in GUS expression was observed over time following *Agrobacterium* infiltration (Figure A3.2A-F). With one exception (*HSP101B*, no temperature treatment), all experiments, regardless of construct, showed the lowest level of GUS-positive seedlings at the 72-hour time point, while the greatest number of GUS-positives were observed at 120 hours post-infiltration (Figure A3.2A-F).

Increase in numbers of GUS-positive seedlings over time varied depending on construct. *AtRPL23aA*- and *HSP101B* (cold-treated)-infiltrated seedlings showed a substantial increase in GUS-positive seedlings between the 72 h and 120 h time points: approximately 7-fold and 17-fold increases, respectively (Figure A3.2A, D). In contrast, GUS-positives in heat-treated *HSP101B*-infiltrated seedlings only doubled between 72 and 120 h (Figure A3.2C), and *pCAMBIA*-infiltrated seedlings only showed a 1.3-fold increase in GUS-positive seedlings from 72 to 120 h (Figure A3.2E). Seedlings infiltrated with *HSP101B* (no temperature treatment) showed 2.7 times more GUS positives at 120 h than at 96 h (Figure A3.2B).

Variation in numbers of GUS-positives between constructs

Numbers of GUS-positive seedlings varied in *HSP101B*-infiltrated plants depending on the temperature induction treatment. Seedlings infiltrated with the *HSP101B* construct showed low levels of GUS activity in the absence of an inductive stimulus or under heat treatment, with approximately 1% of seedlings staining for GUS activity 120 h post-infiltration (Figure A3.2F). Cold-treated *HSP101B* seedlings showed over five and six times more GUS-positives than heat-induced or non-induced *HSP101B* seedlings 120 h post-infiltration.

The number of GUS-positive seedlings varied not only between time points and treatments for the same construct but also varied between constructs (Figure A3.2F, Figure A3.3). *AtRPL23aA*- and *HSP101B*-infiltrated seedlings showed generally low numbers of GUS-positive seedlings; at 120 h post-infiltration these seedlings showed a maximum of approximately 4% GUS positives. *pCAMBIA2301*-infiltrated seedlings showed over six times more GUS positives at 120 h following infiltration than the second-highest GUS expressing seedlings (cold-treated *HSP101B*; Figure A3.3).

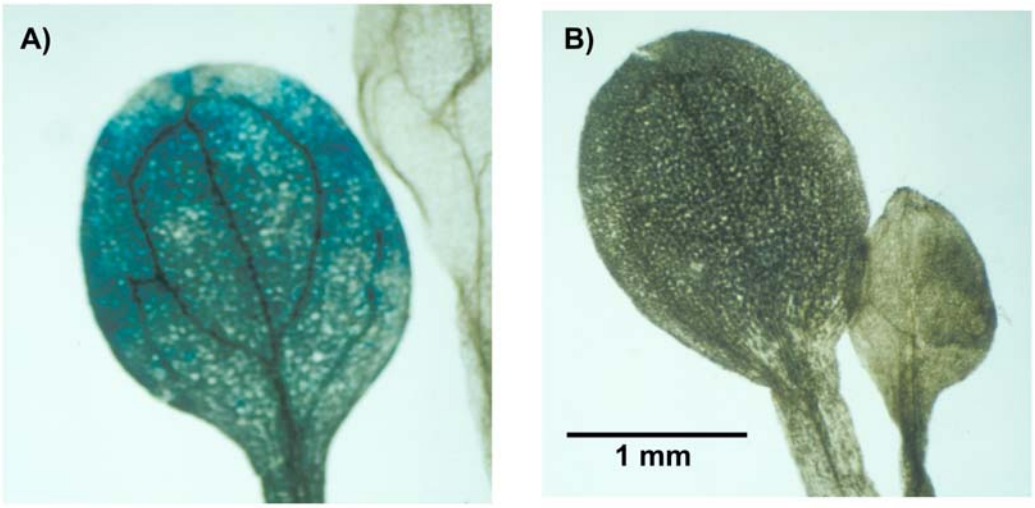
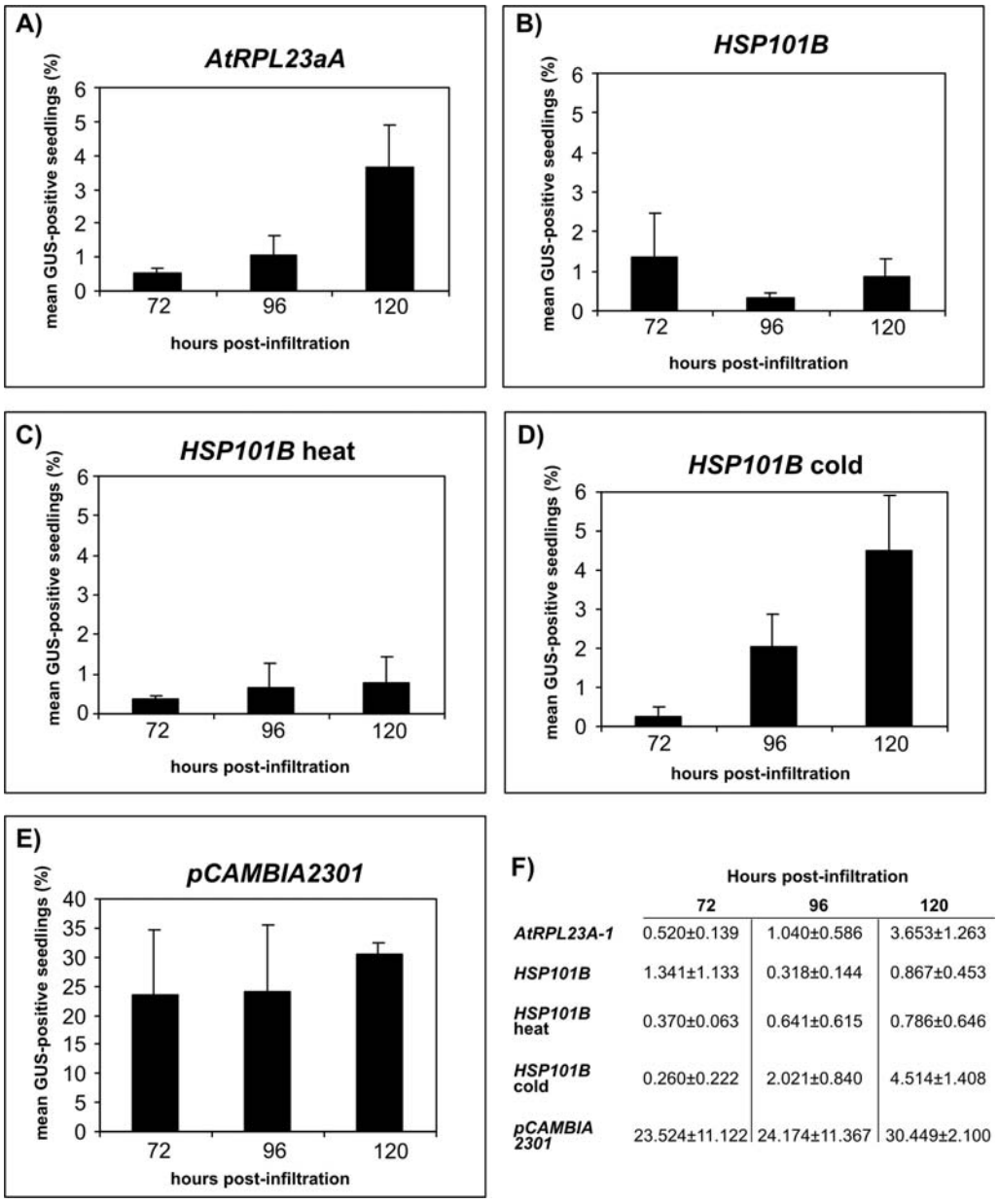


Figure A3.1. Arabidopsis seedlings stained for GUS activity. Blue color indicates presence of GUS enzyme activity. A) *pCAMBIA2301*-infiltrated seedling, GUS positive cotyledon (left) and GUS-negative cotyledon (right), B) vacuum-only negative control seedling, cotyledon and leaf. Photographs at same magnification.



	Hours post-infiltration		
	72	96	120
AtRPL23A-1	0.520±0.139	1.040±0.586	3.653±1.263
HSP101B	1.341±1.133	0.318±0.144	0.867±0.453
HSP101B heat	0.370±0.063	0.641±0.615	0.786±0.646
HSP101B cold	0.260±0.222	2.021±0.840	4.514±1.408
pCAMBIA 2301	23.524±11.122	24.174±11.367	30.449±2.100

Figure A3.2. Percentage of seedlings positive for GUS activity at each of three time points. a) *AtRPL23aA*-infiltrated seedlings, b) *HSP101B*-infiltrated seedlings, c) *HSP101B*-infiltrated seedlings, heat-treated, d) *HSP101B*-infiltrated seedlings, cold-treated, e) *pCAMBIA2301*-infiltrated seedlings, f) GUS positives per time point per construct. Number of GUS positive seedlings are expressed as mean percentages ± SE (n=3).

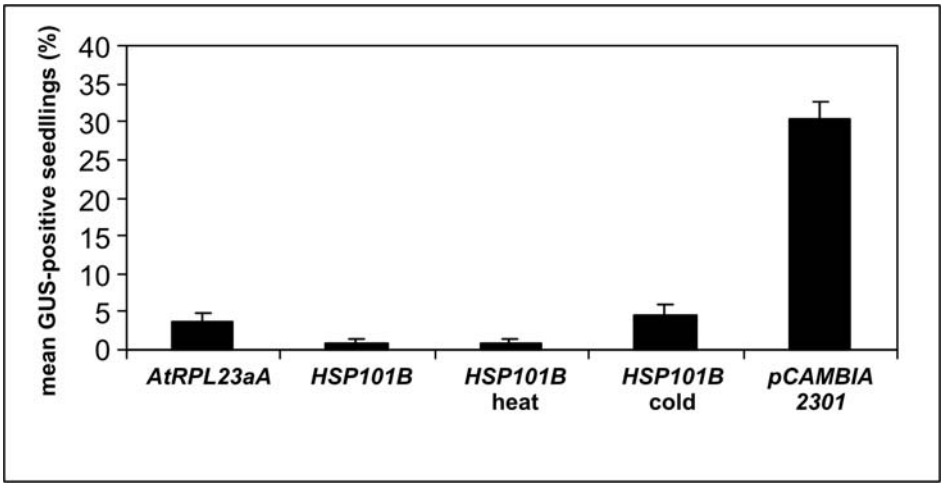


Figure A3.3. Comparison of GUS-positive seedlings at 120 hours post-infiltration for all constructs tested.

Discussion

We have developed a transient assay system for Arabidopsis that is rapid, efficient, economical, and utilizes standard techniques without specialized equipment. Less than three weeks are required to perform the assay in its entirety, from plating the seed to staining the seedlings, and multiple regulatory regions (including both viral, e.g. *CaMV 35S*, and plant regulatory regions) can be analyzed for activity. The use of an *Agrobacterium*-mediated method requires no special equipment, the infiltration medium is simple, and the transient nature of the assay bypasses the problems associated with stable transformation, i.e. silencing due to position effects or transgene copy number. This system is ideal for testing reporter gene constructs to verify activity prior to undergoing stable transformation, which can entail months of work only to discover no reporter gene activity in the resulting transgenics.

The increase in reporter gene activity over five days following *Agrobacterium* infiltration is similar to results reported for transient assays using mature plants (Rakouský et al., 1997). It has been suggested that this may be due to an increase in stable transformation over time or the stability of GUS itself (Rakouský et al., 1997). GUS is a stable enzyme, enabling the protein to accumulate over time (Jefferson, 1987; Martin et al., 1992). However, this does not fully explain why seedlings infiltrated with certain constructs showed a greater increase in GUS-positives over time than seedlings infiltrated with other constructs. Seedlings transformed with the strong constitutive regulator, the *CaMV 35S* promoter present in the *pCAMBIA* construct, only showed a slight increase (1.3-fold) in GUS-positives between 72 and 120 h. In contrast, seedlings infiltrated with the *AtRPL23aA* regulatory region or the *HSP101B* regulatory region and subsequently cold treated, showed greater increases in GUS-positives over time (7- and 17-fold increases over 120 h respectively) than the *pCAMBIA2301*-infiltrated seedlings. GUS-positive numbers for the lowest (*HSP101B*, no induction; *HSP101B*, heat treated) and highest (*pCAMBIA*) expressers remained much more constant than the other constructs tested in this system. The small increase in GUS-positives over time in seedlings infiltrated with low-expressing constructs may simply be due to accumulation of GUS above a threshold over time. Differential increases in GUS-positive seedlings

over time following infiltration with different constructs may reflect a requirement for specific transcription factors for activation of the various transgenic regulatory elements.

The variability in the assay system, as reflected by the standard error for each time point, may be due to the age of plants used. Young, partially expanded tobacco leaves have previously been reported to show much greater variability in GUS expression than more mature leaves (Yang et al., 2000). The advantage of the Arabidopsis seedling assay is that the number of seedlings that can be screened at any one time is very high, with an average of 1502 seedlings per time point per construct sampled using this protocol. The high number of seedlings screened negates any seedling age differences and increases the probability of observing GUS-positives, even when seedlings are transformed with reporter genes controlled by weak regulatory elements.

In the assay presented here, a "GUS-positive" was counted as any seedling showing staining for GUS activity, unlike other assays (Rossi et al., 1993) where blue spots on each GUS-positive seedling were counted as individual transformation events. Our approach will underestimate the number of transformation events; even if an attempt is made to count the number of blue spots on positive seedlings it is difficult to estimate these counts when transformation efficiency is very high (Rossi et al., 1993). Variability in amounts and intensity of staining is inherent with a histochemical assay; however, the Arabidopsis seedling transient assay has not been developed as a quantitative tool. This assay was developed to examine the functionality of constructs *in planta* and to examine relative regulatory region strength and responsiveness to inductive stimuli.

Although the strong constitutive *CaMV 35S* promoter driving GUS expression in *pCAMBIA2301* resulted in high levels of GUS expression, endogenous plant promoters were also found to be active in the transient assay system. Regulatory region strength was reflected in the percentage of GUS-positive seedlings for a variety of constructs; the Arabidopsis *AtRPL23aA* and *HSP101B* regulatory regions resulted in the lowest number of GUS-positives in the absence of inductive stimuli. *HSP101B* has been shown to drive low-level constitutive expression in stably transformed flax (Cross, 2002) and canola (Young, 2003). Three intact HSEs are required for high levels of transcription in HSP

promoters (Barros et al., 1992; Schöffl et al., 1998) therefore a low level of *HSP101B* regulatory region (containing two HSEs) activity under heat stress was not unexpected.

In addition to its utility in analyzing promoters and other gene regulatory regions the Arabidopsis seedling assay is a quick and easy method for testing reporter gene constructs prior to stable transformation. The constructs can subsequently be used for stable transformation. The system could also be used for applications such as *in planta* protein overexpression.

REFERENCES

Adilakshmi, T., and Laine, R.O. (2002) Ribosomal protein S25 mRNA partners with MTF-1 and La to provide a p53-mediated mechanism for survival or death. J. Biol. Chem. 277: 4147-4151.

Agatep, R., Kirkpatrick, R.D., Parchaliuk, D.L., Woods, R.A., and Gietz, R.D. (1998) Transformation of *Saccharomyces cerevisiae* by the lithium acetate/single-stranded carrier DNA/polyethylene glycol (LiAc/ssDNA/PEG) protocol. Technical Tips Online (http://tto.trends.com).

Agrawal, M.G., and Bowman, L.H. (1987) Transcriptional and translational regulation of ribosomal protein formation during mouse myoblast differentiation. J. Biol. Chem. 262: 4868-4875.

Agrawal, R.K., Penczek, P., Grassucci, R.A., and Frank, J. (1998) Visualization of elongation factor G on the *Escherichia coli* 70S ribosome: the mechanism of translocation. Proc. Natl. Acad. Sci. U.S.A. 95: 6134-6138.

Agrawal, R.K., Heagle, A.B., Penczek, P., Grassucci, R.A., and Frank, J. (1999) EF-G-dependent GTP hydrolysis induces translocation accompanied by large conformational changes in the 70S ribosome. Nat. Struct. Biol. 6: 643-647.

Allen, T., Shen, P., Samsel, L., Liu, R., Lindahl, L., and Zengel, J.M. (1999) Phylogenetic analysis of L4-mediated autogenous control of the S10 ribosomal protein operon. J. Bacteriol. 181: 6124-6132.

Allen, T.D., Watkins, T., Lindahl, L., and Zengel, J.M. (2004) Regulation of ribosomal protein synthesis in *Vibrio cholerae*. J. Bacteriol. 186: 5933-5937.

Allers, J., and Shamoo, Y. (2001) Structure-based analysis of protein-RNA interactions using the program ENTANGLE. J. Mol. Biol. 311: 75-86.

Aloni, R., Peleg, D., and Meyuhas, O. (1992) Selective translational control and nonspecific posttranscriptional regulation of ribosomal protein gene expression during development and regeneration of rat liver. Mol. Cell. Biol. 12: 2203-2212.

Amaldi, F., Bozzoni, I., Beccari, E., and Pierandrei-Amaldi, P. (1989) Expression of ribosomal protein genes and regulation of ribosome biosynthesis in *Xenopus* development. Trends Biochem. Sci. 14: 175-178.

Amaldi, F., Camacho-Vanegas, O., Cardinali, B., Cecconi, F., Crosio, C., Loreni, F,

Mariottini, P., Pellizzoni, L., and Pierandrei-Amaldi, P. (1995) Structure and expression of ribosomal protein genes in *Xenopus laevis*. Biochem. Cell Biol. 73: 969-977.

Andersson, S., Sæbøe-Larssen, S., Lambertsson, A., Merriam, J., and Jacobs-Lorena, M. (1994) A Drosophila third chromosome *Minute* locus encodes a ribosomal protein. Genetics 137: 513-520.

Arabi, A., Wu, S., Ridderstråle, K., Bierhoff, H., Shiue, C., Fatyol, K., Fahlén, S., Hydbring, P., Söderberg, O., Grummt, I., Larsson, L.-G., and Wright, A.P.H. (2005) c-Myc associates with ribosomal DNA and activates RNA polymerase I transcription. Nat. Cell Biol. 7: 303-310.

Arabidopsis Genome Initiative, AGI (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796-815.

Arnold, R.J., and Reilly, J.P. (1999) Observation of *Escherichia coli* ribosomal proteins and their posttranslational modifications by mass spectrometry. Anal. Biochem. 269: 105-112.

Bacher, G., Lütcke, H., Jungnickel, B., Rapoport, T.A., and Dobberstein, B. (1996) Regulation by the ribosome of the GTPase of the signal-recognition particle during protein targeting. Nature 381: 248-251.

Bacher, G., Pool, M., and Dobberstein, B. (1999) The ribosome regulates the GTPase of the β-subunit of the signal recognition particle receptor. J. Cell Biol. 146: 723-730.

Bailey-Serres, J. (1998) Cytoplasmic ribosomes of higher plants. In: A look beyond transcription: Mechanisms determining mRNA stability and translation in plants. Edited by: Bailey-Serres, J., and Gallie, D.R. American Association of Plant Physiologists, Rockwell, MD. pp: 125-144.

Bailey-Serres, J., and Freeling, M. (1990) Hypoxic stress-induced changes in ribosomes of maize seedling roots. Plant Physiol. 94: 1237-1243.

Baker, S.S., Wilhelm, K.S., and Thomashow, M.F. (1994) The 5'-region of *Arabidopsis thaliana cor15a* has *cis*-acting elements that confer cold-, drought-, and ABA-regulated gene expression. Plant Mol. Biol. 24:701-713.

Ban, N., Nissen, P., Hansen, J., Capel, M., Moore, P.B., and Steitz, T.A. (1999)

Placement of protein and RNA structures into a 5Å-resolution map of the 50S ribosomal subunit. Nature 400: 841-847.

Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. Science 289: 905-920.

Barakat, A., Szick-Miranda, K., Chang, I.-F., Guyot, R., Blanc, G., Cooke, R., Delseny, M., and Bailey-Serres, J. (2001) The organization of cytoplasmic ribosomal protein genes in the Arabidopsis genome. Plant Physiol. 127: 398-415.

Barros, M.D., Czarnecka, E., and Gurley, W.B. (1992) Mutational analysis of a plant heat shock element. Plant Mol. Biol. 19:665-675.

Bate, N., and Twell, D. (1998) Function architecture of a late pollen promoter: multiple stage-specific and co-dependent activator elements. Plant Mol. Biol. 37: 859-869.

Bayfield, M.A., Dahlberg, A.E., Schulmeister, U., Dorner, S., and Barta, A. (2001) A conformational change in the ribosomal peptidyl transferase center upon active/inactive transition. Proc. Natl. Acad. Sci. U.S.A. 98: 10096-10101.

Beckmann, R., Bubeck, D., Grassucci, R., Penczek, P., Verschoor, A., Blobel, G., and Frank, J. (1997) Alignment of conduits for the nascent polypeptide chain in the ribosome-Sec61 complex. Science 278: 2123-2126.

Beckmann, R., Spahn, C.M., Eswar, N., Helmers, J., Penczek, P.A., Sali, A, Frank, J., and Blobel, G. (2001) Architecture of the protein-conducting channel associated with the translating 80S ribosome. Cell 107: 361-372.

Beltrán-Peña, E., Aguilar, R., Ortíz-López, A., Dinkova, T.D., and Sánchez de Jiménez, E. (2002) Auxin stimulates S6 ribosomal protein phosphorylation in maize thereby affecting protein synthesis regulation. Physiol. Plant. 115: 291-297.

Beltrán-Peña, E., Ortíz-López, A., and Sánchez de Jiménez, E. (1995) Synthesis of ribosomal proteins from stored mRNAs early in seed germination. Plant Mol. Biol. 28: 327-336.

Berberich, T., Uebeler, M., and Feierabend, J. (2000) cDNA cloning of cytoplasmic ribosomal protein S7 of winter rye (*Secale cereale*) and its expression in low-temperature-treated leaves. Biochim. Biophys. Acta 1492: 276-279.

Berisio, R., Schluenzen, F., Harms, J., Bashan, A., Auerbach, T., Baram, D., and Yonath, A. (2003) Structural insight into the role of the ribosomal tunnel in cellular regulation. Nat. Struct. Biol. 10: 366-370.

Bielka, H. (1982) The Eukaryotic Ribosome. Berlin: Springer-Verlag.

Bixby, J.A., and Brown, G.N. (1975) Ribosomal changes during induction of cold hardiness in black locust seedlings. Plant Physiol. 56: 617-621.

Blaha, G., Wilson, D.N., Stoller, G., Fischer, G., Willumeit, R., and Nierhaus, K.H. (2003) Localization of the trigger factor binding site on the ribosomal 50S subunit. J. Mol. Biol. 326: 887-897.

Bloemink, M.J., and Moore, P.B. (1999) Phosphorylation of ribosomal protein L18 is required for its folding and binding to 5S rRNA. Biochemistry 38: 13385-13390.

Bollen, G.H.P.M., Molenaar, C.M.T., Cohen, L.H., van Raamsdonk-Duin, M.M.C., Mager, W.H., and Planta, R.J. (1982) Ribosomal protein genes of yeast contain intervening sequences. Gene 18: 29-37.

Bonham-Smith, P.C., Oancia, T.L., and Moloney, M.M. (1992) Cytoplasmic ribosomal protein S15a from *Brassica napus*: molecular cloning and developmental expression in mitotically active tissues. Plant Mol. Biol. 18: 909-919.

Boublik, M., Mandiyan, V., and Tumminia, S. (1990) Potential of electron microscopic techniques for structural analysis of ribosomes. In: The ribosome: structure, function, and evolution. Edited by: Hill, W.E., Dahlberg, A., Garrett, R.A., Moore, P.B., Schlessinger, D., and Warner, J.R. American Society for Microbiology, Washington, D.C. pp: 114-122.

Bourdon, V., Harvey, A., and Lonsdale, D.M. (2001) Introns and their positions affect the translational activity of mRNA in plant cells. EMBO Rep. 2: 394-398.

Brodersen, D.E., Clemons, W.M.Jr., Carter, A.P., Wimberly, B.T., and Ramakrishnan, V. (2002) Crystal structure of the 30S ribosomal subunit from *Thermus thermophilus*: structure of the proteins and their interactions with 16S RNA. J. Mol. Biol. 316: 725-768.

Browning, K.S. (1996) The plant translational apparatus. Plant Mol. Biol. 32: 107-144.

Buisson, M., and Reboud, A.M. (1982) Carbodiimide-induced protein-RNA crosslinking in mammalian ribosomal subunits. FEBS Lett. 148: 247-250.

Buskiewicz, I., Deuerling, E., Gu, S.-Q., Jöckel, J., Rodnina, M.V., Bukau, B., and Wintermeyer, W. (2004) Trigger factor binds to ribosome-signal-recognition particle (SRP) complexes and is excluded by binding of the SRP receptor. Proc. Natl. Acad. Sci. U.S.A. 101: 7902-7906.

Campbell, J.H. (1991) An RNA replisome as the ancestor of the ribosome. J. Mol. Evol. 32: 3-5.

Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D'Souza, L.M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., Muller, K.M., Pande, N., Shang, Z., Yu, N., and Gutell, R.R. (2002) The Comparative RNA Web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BioMed Central Bioinformatics 3: 2.(Correction in BioMed Central Bioinformatics 3:15).

Carter, A.P., Clemons, W.M., Brodersen, D.E., Morgan-Warren, R.J., Wimberly, B.T., and Ramakrishnan, V. (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. Nature 407: 340-348.

Carrari, F., Fernie, A.R., and Iusem, N.D. (2004) Heard it through the grapevine? ABA and sugar cross-talk: the ASR story. Trends Plant Sci. 9: 57-59.

Casati, P., and Walbot, V. (2003) Gene expression profiling in response to ultraviolet radiation in maize genotypes with varying flavonoid content. Plant Physiol. 132: 1739-1754.

Caspersson, T., and Schultz, J. (1939) Pentose nucleotides in the cytoplasm of growing tissues. Nature 143: 602-603.

Caspersson, T., and Schultz, J. (1940) Ribonucleic acids in both nucleus and cytoplasm, and the function of the nucleolus. Proc. Natl. Acad. Sci. U.S.A. 26: 507-515.

Cate, J.H., Yusupov, M.M., Yusupova, G.Z., Earnest, T.N., and Noller, H.F. (1999) X-ray crystal structures of 70S ribosome functional complexes. Science 285: 2095-2104.

Cech, T.R., and Bass, B.L. (1986) Biological catalysis by RNA. Ann. Rev. Biochem. 55: 599-629.

Cercós, M., Gómez-Cadenas, A., and Ho, T-H. D. (1999) Hormonal regulation of a

cysteine proteinase gene, *EPB-1*, in barley aleurone layers: cis- and trans-acting elements involved in the co-ordinated gene expression regulated by gibberellins and abscisic acid. Plant J. 19: 107-118.

Chang, I.-F., Szick-Miranda, K., Pan, S., and Bailey-Serres, J. (2005) Proteomic characterization of evolutionarily conserved and variable proteins of Arabidopsis cytosolic ribosomes. Plant Physiol. 137: 848-862.

Charollais, J., Dreyfus, M., and Iost, I. (2004) CsdA, a cold-shock RNA helicase from *Escherichia coli*, is involved in the biogenesis of 50S ribosomal subunit. Nuc. Acids Res. 32: 2751-2759.

Chavatte, L., Brown, B.A.II, and Driscoll, D.M. (2005) Ribosomal protein L30 is a component of the UGA-selenocysteine recoding machinery in eukaryotes. Nat. Struct. Mol. Biol. 12: 408-416.

Chenuil, A., Solignac, M., and Bernard, M. (1997) Evolution of the large-subunit ribosomal RNA binding site for protein L23/25. Mol. Biol. Evol. 14: 578-588.

Cheong, Y.H., Chang, H-S., Gupta, R., Wang, X., Zhu, T., and Luan, S. (2002) Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in Arabidopsis. Plant Physiol. 129: 661-677.

Cherepneva, G.N., Schmidt, K.-H., Kulaeva, O.N., Oelmuller, R., and Kusnetsov, V.V. (2003) Expression of the ribosomal proteins S14, S16, L13a, and L30 is regulated by cytokinin and abscisic acid. Implication of the involvement of phytohormones in translational processes. Plant Sci. 165: 925-932.

Chevalier, C., Le Querrec, F., and Raymond, P. (1996) Sugar levels regulate the expression of ribosomal protein genes encoding protein S28 and ubiquitin-fused protein S27a in maize primary root tips. Plant Sci. 117: 95-105.

Chinnusamy, V., Schumaker, K., and Zhu, J.K. (2004) Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants. J. Expt. Bot. 55: 225-236.

Chung, H., Hong, D.-P., Jung, J.-Y., Kim, H.-J., Jang, K.-S., Sheen, Y.-Y., Ahn, J.-I., Lee, Y.-S., and Kong, G. (2005) Comprehensive analysis of differential gene expression profiles on carbon tetrachloride-induced rat liver injury and

regeneration. Toxicol. Appl. Pharmacol. 206: 27-42.

Chung, S., and Perry, R.P. (1989) Importance of introns for expression of mouse ribosomal protein gene rpL32. Mol. Cell Biol. 9: 2075-2082.

Chung, S., and Perry, R.P. (1993) The importance of downstream δ-factor binding elements for the activity of the rpL32 promoter. Nuc. Acids Res. 21: 3301-3308.

Claude, A. (1941) Particulate components of cytoplasm. Cold Spring Harbor Symp. Quant. Biol. 9: 263-271.

Claude, A. (1943) The constitution of protoplasm. Science 97: 451-456.

Clemons, W.M. Jr., May, J.L.C., Wimberly, B.T., McCutcheon, J.P., Capel, M.S., and Ramakrishnan, V. (1999) Structure of a bacterial 30S ribosomal subunit at 5.5 Å resolution. Nature 400: 833-840.

Clough, S.J., and Bent, A.F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. 16: 735-743.

Coenye, T., and Vandamme, P. (2005) Organisation of the *S10*, *spc* and *alpha* ribosomal protein gene clusters in prokaryotic genomes. FEMS Micro. Lett. 242: 117-126.

Cooperman, B.S., Wooten, T., Romero, D.P., and Traut, R.R. (1995) Histidine 229 in protein L2 is apparently essential for 50S peptidyl transferase activity. Biochem. Cell Biol. 73: 1087-1094.

Crick, F.H.C. (1968) The origin of the genetic code. J. Mol. Biol. 38: 367-379.

Cross, R.H. (2002) Heat stress effects on flowering and reproduction in *Linum usitatissimum* (flax). MSc thesis. University of Saskatchewan, Saskatoon, SK.

Cukras, A.R., and Green, R. (2005) Multiple effects of S13 in modulating the strength of intersubunit interactions in the ribosome during translation. J. Mol. Biol. 349: 47-59.

Culver, G.M. (2003) Assembly of the 30S ribosomal subunit. Biopolymers 68: 234-249.

Curie, C., Liboz, T., Bardet, C., Gander, E., Médale, C., Axelos, M., and Lescure, B. (1991) *Cis* and *trans*-acting elements involved in the activation of *Arabidopsis thaliana* A1 gene encoding the translation elongation factor EF-1a. Nuc. Acids Res. 19: 1305-1310.

Dai, Z., Gao, J., An, K., Lee, J.M., Edwards, G.E., and An, G. (1996) Promoter elements

controlling developmental and environmental regulation of a tobacco ribosomal protein gene *L34*. Plant Mol. Biol. 32: 1055-1065.

Darbinian, N., Gallia, G.L., and Khalili, K. (2001) Helix-destabilizing properties of the human single-stranded DNA- and RNA-binding protein Purα. J. Cell. Biochem. 80: 589-595.

De Angelis, M., Di Cagno, R., Huet, C., Crecchio,C., Fox, P.F., and Gobbetti, M. (2004) Heat shock response in *Lactobacillus plantarum*. App. Environ. Micro. 70: 1336-1346.

Dez, C., and Tollervey, D. (2004) Ribosome synthesis meets the cell cycle. Curr. Opin. Microbiol. 7: 631-637.

Dick, F.A., and Trumpower, B.L. (1998) Heterologous complementation reveals that mutant alleles of *QSR1* render 60S ribosomal subunits unstable and translationally inactive. Nuc. Acids Res. 26: 2442-2448.

Diedrich, G., Spahn, C.M.T., Stelzl, U., Schäfer, M.A., Wooten, T., Bochkariov, D.E., Cooperman, B.S., Traut, R.R., and Nierhaus, K.H. (2000) Ribosomal protein L2 is involved in the association of the ribosomal subunits, tRNA binding to A and P sites and peptidyl transfer. EMBO J. 19: 5241-5250.

Draper, D.E., and Reynaldo, L.P. (1999) RNA binding strategies of ribosomal proteins. Nuc. Acids Res. 27: 381-388.

Dresselhaus, T., Cordts, S., Heuer, S., Sauter, M., Lörz, H., and Kranz, E. (1999) Novel ribosomal genes from maize are differentially expressed in the zygotic and somatic cell cycles. Mol. Gen. Genet. 261: 416-427.

Dube, P., Wieske, M., Stark, H., Schatz, M., Stahl, J., Zemlin, F., Lutsch, G., and van Heel, M. (1998a) The 80S rat liver ribosome at 25Å resolution by electron cryomicroscopy and angular reconstitution. Structure 6: 389-399.

Dube, P., Bacher, G., Stark, H., Mueller, F., Zemlin, F., van Heel, M., and Brimacombe, R. (1998b) Correlation of the expansion segments in mammalian rRNA with the fine structure of the 80S ribosome; a cryoelectron microscopic reconstruction of the rabbit reticulocyte ribosome at 21 Å resolution. J. Mol. Biol. 279: 403-421.

Egebjerg, J., Christiansen, J., and Garrett, R.A. (1991) Attachment sites of primary binding proteins L1, L2, and L23 on 23S ribosomal RNA of *Escherichia coli*. J

Mol. Biol. 222: 251-264.

El-Baradi, T.T.A.L., de Regt, V.C.H.F., Planta, R.J., Nierhaus, K.H., and Raué, H.A. (1987) Interaction of ribosomal proteins L25 from yeast and EL23 from *E. coli* with yeast 26S and mouse 28S rRNA. Biochimie 69: 939-948.

El-Baradi, T.T.A.L., Raué, H.A., de Regt, V.C.H.F., and Planta, R.J. (1984) Stepwise dissociation of yeast 60S ribosomal subunits by LiCl and identification of L25 as a primary 26S rRNA binding protein. Eur. J. Biochem. 144: 393-400.

El-Baradi, T.T.A.L., Raué, H.A., de Regt, V.C.H.F., Verbree, E.C., and Planta, R.J. (1985) Yeast ribosomal protein L25 binds to an evolutionary conserved site on yeast 26S and *E. coli* 23S rRNA. EMBO J. 4: 2101-2107.

Elcock, A.H., and McCammon, J.A. (1998) Electrostatic contributions to the stability of halophilic proteins. J. Mol. Biol. 280: 731-748.

Elmayan, T., and Tepfer, M. (1995) Evaluation in tobacco of the organ specificity and strength of the rol D promoter, domain A of the 35S promoter and the $35S^2$ promoter. Transgenic Res. 4: 388-396.

Fan, W., Christensen, M., Eichler, E., Zhang, X., and Lennon, G. (1997) Cloning, sequencing, gene organization, and localization of the human ribosomal protein RPL23A gene. Genomics 46: 234-239.

Fatica, A., and Tollervey, D. (2002) Making ribosomes. Curr. Opin. Cell Biol. 14: 313-318.

Fehling, E., and Weidner, M. (1986) Temperature characteristics and adaptive potential of wheat ribosomes. Plant Physiol. 80: 181-186.

Fehling, E., and Weidner, M. (1988) Adaptive potential of wheat ribosomes toward heat depends on the large ribosomal subunit and ribosomal protein phosphorylation. Plant Physiol. 87: 562-565.

Fennoy, S.L., and Bailey-Serres, J. (1995) Post-transcriptional regulation of gene expression in oxygen-deprived roots of maize. Plant J. 7: 287-295.

Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004) Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. Nature 431: 590-596.

Fleming, G., Belhumeur, P., Skup, D., and Fried, H.M. (1989) Functional substitution of

mouse ribosomal protein L27' for yeast ribosomal protein L29 in yeast ribosomes. Proc. Natl. Acad. Sci. U.S.A. 86: 217-221.

Frank, J., and Agrawal, R.K. (2000) A ratchet-like inter-subunit reorganization of the ribosome during translocation. Nature 406: 318-322.

Fromont-Racine, M., Senger, B., Saveanu, C., and Fasiolo, F. (2003) Ribosome assembly in eukaryotes. Gene 313: 17-42.

Fulga, T.A., Sinning, I., Dobberstein, B., and Pool, M.R. (2001) SRβ coordinates signal sequence release from SRP with ribosome binding to the translocon. EMBO J. 20: 2338-2347.

Gabashvili, I.S., Agrawal, R.K., Spahn, C.M.T., Grassucci, R.A., Svergun, D.I., Frank, J., and Penczek, P. (2000) Solution structure of the *E. coli* 70S ribosome at 11.5 Å resolution. Cell 100: 537-549.

Gantt, J.S., and Key, J.L. (1983) Auxin-induced changes in the level of translatable ribosomal protein messenger ribonucleic acids in soybean hypocotyl. Biochemistry 22: 4131-4139.

Gantt, J.S., and Key, J.L. (1985) Coordinate expression of ribosomal protein mRNAs following auxin treatment of soybean hypocotyls. J. Biol. Chem. 260: 6175-6181.

Gao, J., Kim, S.-R., Chung, Y.-Y., Lee, J.M., and An, G. (1994) Developmental and environmental regulation of two ribosomal protein genes in tobacco. Plant Mol. Biol. 25: 761-770.

Gao, H., Sengupta, J., Valle, M., Korostelev, A., Eswar, N., Stagg, S.M., Van Roey, P., Agrawal, R.K., Harvey, S.C., Sali, A., Chapman, M.S., and Frank, J. (2003) Study of the structural dynamics of the *E. coli* 70S ribosome using real-space refinement. Cell 113: 789-801.

García-Hernández , M., Davies, E., and Staswick, P.E. (1994) *Arabidopsis* p40 homologue. J. Biol. Chem. 269: 20744-20749.

Genuario, R.R., and Perry, R.P. (1996) The GA-binding protein can serve as both an activator and repressor of *ribosomal protein* gene transcription. J. Biol. Chem. 271: 4388-4395.

Geyer, P.K., Meyuhas, O., Perry, R.P., and Johnson, L.F. (1982) Regulation of

ribosomal protein mRNA content and translation in growth-stimulated mouse fibroblasts. Mol. Cell. Biol. 2: 685-693.

Giannino, D., Frugis, G., Ticconi, C., Florio, S., Mele, G., Santini, L., Cozza, R., Bitonti, M.B., Innocenti, A., and Mariotti, D. (2000) Isolation and molecular characterisation of the gene encoding the cytoplasmic ribosomal protein S28 in *Prunus persica* [L.] Batsch. Mol. Gen. Genet. 263: 201-212.

Giavalisco, P., Wilson, D., Kreitler, T., Lehrach, H., Klose, J., Gobom, J., and Fucini, P. (2005) High heterogeneity within the ribosomal proteins of the *Arabidopsis thaliana* 80S ribosome. Plant Mol. Biol. 57: 577-591.

Giraudat, J., Parcy, F., Bertauche, N., Gosti, F., Leung, J., Morris, P-C., Bouvier-Durand, M., and Vartanian, N. (1994) Current advances in abscisic acid action and signaling. Plant Mol.Biol. 26: 1557-1577.

Gomez-Lorenzo, M.G., Spahn, C.M.T., Agrawal, R.K., Grassucci, R.A., Penczek, P., Chakraburtty, K., Ballesta, J.P.G., Lavandera, J.L., Garcia-Bustos, J.F., and Frank, J. (2000) Three-dimensional cryo-electron microscopy localization of EF2 in the *Saccharomyces cerevisiae* 80S ribosome at 17.5 Å resolution. EMBO J. 19: 2710-2718.

Grandori, C., Gomez-Roman, N., Felton-Edkins, Z.A., Ngouenet, C., Galloway, D.A., Eisenman, R.N., and White, R.J. (2005) c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I. Nat. Cell Biol. 7: 311-318.

Granneman, S., and Baserga, S.J. (2004) Ribosome biogenesis: of knobs and RNA processing. Exp. Cell Res. 296: 43-50.

Grewal, S.S., Li, L., Orian, A., Eisenman, R.N., and Edgar, B.A. (2005) Myc-dependent regulation of ribosomal RNA synthesis during Drosophila development. Nat. Cell Biol. 7: 295-302.

Grossman, A.R. (2005) Paths toward algal genomics. Plant Physiol. 137: 410-427.

Gu, S.-Q., Peske, F., Wieden, H.-J., Rodnina, M.V., and Wintermeyer, W. (2003) The signal recognition particle binds to protein L23 at the peptide exit of the *Escherichia coli* ribosome. RNA 9: 566-573.

Gualerzi, C., Janda, H.G., Passow, H., and Stöffler, G. (1974) Studies on the protein

moiety of plant ribosomes. J. Biol. Chem. 249: 3347-3355.

Guilfoyle, T., Hagen, G., Ulmasov, T., and Murfett, J. (1998) How does auxin turn on genes? Plant Physiol. 118: 341-347.

Hadjiolov, A.A. (1985) The nucleolus and ribosome biogenesis. Springer-Verlag, New York.

Halic, M., Becker, T., Pool, M.R., Spahn, C.M.T., Grassucci, R.A., Frank, J., and Beckmann, R. (2004) Structure of the signal recognition particle interacting with the elongation-arrested ribosome. Nature 427: 808-814.

Halic, M., Becker, T., Frank, J., Spahn, C.M.T., and Beckmann, R. (2005) Localization and dynamic behavior of ribosomal protein L30e. Nat. Struct. Mol. Biol. 12: 467-468.

Hampl, H., Schulze, H., and Nierhaus, K.H. (1981) Ribosomal components from *Escherichia coli* 50S subunits involved in the reconstitution of peptidyl transferase activity. J. Biol. Chem. 256: 2284-2288.

Hansen, J.L., Schmeing, T.M., Moore, P.B., and Steitz, T.A. (2002) Structural insights into peptide bond formation. Proc. Natl. Acad. Sci. U.S.A. 99: 11670-11675.

Hanson, C.L., Videler, H., Santos, C., Ballesta, J.P., and Robinson, C.V. (2004) Mass spectrometry of ribosomes from *Saccharomyces cerevisiae*: implications for assembly of the stalk complex. J. Biol. Chem. 279: 42750-42757.

Hardy, S.J.S (1975) The stoichiometry of the ribosomal proteins of *Escherichia coli*. Mol. Gen. Genet. 140: 253-274.

Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. (2001) High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. Cell 107: 679-688.

Harrison, P.M., Hegyi, H., Balasubramanian, S., Luscombe, N.M., Bertone, P., Echols, N., Johnson, T., and Gerstein, M. (2002) Molecular fossils in the human genome: identification and analysis of the pseudogenes in chromosomes 21 and 22. Genome Res. 12: 272-280.

Held, W.A., Ballou, B., Mizushima, S., and Nomura, M. (1974) Assembly mapping of 30S ribosomal proteins from *Escherichia coli*: further studies. J. Biol. Chem. 249: 3103-3111.

Herold, M., and Nierhaus, K.H. (1987) Incorporation of six additional proteins to complete the assembly map of the 50S subunit from *Escherichia coli* ribosomes. J. Biol. Chem. 262: 8826-8833.

Higo, K., Ugawa, Y., Iwamoto, M., and Korenaga, T. (1999) Plant cis-acting regulatory DNA elements (PLACE) database:1999. Nuc. Acids Res. 27: 297-300.

Hihara, Y., Umeda, M., Hara, C., Toriyama, K., and Uchimiya, H. (1994) Nucleotide sequence of a rice acidic ribosomal phosphoprotein P0 cDNA. Plant Physiol. 105: 753-754.

Hoang, L., Fredrick, K., and Noller, H.F. (2004) Creating ribosomes with an all-RNA 30S subunit P site. Proc. Nat. Acad. Sci. U.S.A. 101: 12439-12443.

Hoekema, A., Hirsch, P.R., Hooykaas, P.J.J., and Schilperoort, R.A. (1983) A binary plant vector strategy based on separation of *vir*- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. Nature 303:179-180.

Hong, S., and Vierling, E. (2001) HSP101 is necessary for heat tolerance but dispensable for development and germination in the absence of stress. Plant J. 27: 25-35.

Hughes, A.L., and Friedman, R. (2005) Expression patterns of duplicate genes in the developing root in *Arabidopsis thaliana*. J. Mol. Evol. 60: 247-256.

Hulm, J.L., McIntosh, K.B., and Bonham-Smith, P.C. (2005) Variation in transcript abundance among the four members of the Arabidopsis thaliana *RIBOSOMAL PROTEIN S15a* gene family. Plant Sci. 169: 267-278.

Jagannanthan, I., and Culver, G.M. (2003) Assembly of the central domain of the 30S ribosomal subunit: roles for the primary binding ribosomal proteins S15 and S8. J. Mol. Biol. 330: 373-383.

Jakel, S., and Gorlich, D. (1998) Importin β, transportin, RanBP5, and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells. EMBO J. 17: 4491-4502.

Jeeninga, R.E., Venema, J., and Raué, H.A. (1996) Rat RL23a ribosomal protein efficiently competes with its *Saccharomyces cerevisiae* L25 homologue for assembly into 60S subunits. J. Mol. Biol. 263: 648-656.

Jefferson, R.A. (1987) Assaying chimeric genes in plants: the GUS gene fusion system. Plant Mol. Biol. Rep. 5:387-405.

Jenni, S., and Ban, N. (2003) The chemistry of protein synthesis and voyage through the ribosomal tunnel. Curr. Opin. Struct. Biol. 13: 212-219.

Jiang, C., Iu, B., and Singh, J. (1996) Requirement of a CCGAC *cis*-acting element for cold induction of the *BN115* gene from winter *Brassica napus*. Plant Mol. Biol. 30: 679-684.

Jiménez, J.L., Mitchell, M.P., and Sgouros, J.G. (2002) Microarray analysis of orthologous genes: conservation of the translational machinery across species at the sequence and expression level. Genome Biol. 4: R4.1-R4.28.

Jin., A., Itahana, K., O'Keefe, K., and Zang, Y. (2004) Inhibition of HDM2 and activation of p53 by ribosomal protein L23. Mol. Cell. Biol. 24: 7669-7680.

Joanin, P., Gigot, C., and Philipps, G. (1993) cDNA nucleotide sequence and expression of a maize cytoplasmic ribosomal protein S13 gene. Plant Mol. Biol. 21: 701-704.

Johnson, A.E. (2004) Functional ramifications of FRET-detected nascent chain folding far inside the membrane-bound ribosome. Biochem. Soc. Trans. 32: 668-672.

Johnson, S.P., and Warner, J.R. (1987) Phosphorylation of the *Saccharomyces cerevisiae* equivalent of ribosomal protein S6 has no detectable effect on growth. Mol. and Cell. Biol. 7: 1338-1345.

Kaltschmidt, E., and Wittmann, H.G. (1970) Ribosomal proteins, XII. Number of proteins in small and large ribosomal subunits of *Escherichia coli* as determined by two-dimensional gel electrophoresis. Proc. Natl. Acad. Sci. U.S.A. 67: 1276-1282.

Katunin, V.I., Muth, G.W., Strobel, S.A., Wintermeyer, W., and Rodnina, M.V. (2002) Important contribution to catalysis of peptide bond formation by a single ionizing group within the ribosome. Mol. Cell 10: 339-346.

Keenan, R.J., Freymann, D.M., Stroud, R.M., and Walter, P. (2001) The signal recognition particle. Ann. Rev. Biochem. 70: 755-775.

Kenmochi, N., Kawaguchi, T., Rozen, S., Davis, E., Goodman, N., Hudson, T.J., Tanaka,T., and Page, D.C. (1998) A map of 75 human ribosomal protein genes. Genome Res. 8: 509-523.

Kenward, M.G., and Roger, J.H. (1997) Small sample inference for fixed effects from

restricted maximum likelihood. Biometrics 53: 983-997.

Key, J.L., Lin, C.Y., Gifford, E.M. Jr., and Dengler, R. (1966) Relation of 2,4-D-induced growth aberrations to changes in nucleic acid metabolism in soybean seedlings. Bot. Gaz. 127: 87-94.

Kim, J.-H., You, K.-R., Kim, I.H., Cho, B.H., Kim, C.-Y., and Kim, D.-G. (2004) Over-expression of the ribosomal protein L36a gene is associated with cellular proliferation in hepatocellular carcinoma. Hepatology 39: 129-138.

Kim, K.-Y., Park, S.-W., Chung, Y.-S., Chung, C.-H., Kim, J.-I., and Lee, J.-H. (2004) Molecular cloning of low-temperature-inducible ribosomal proteins from soybean. J. Exp. Bot. 55: 1153-1155.

Kirillov, S., Wower, J., Hixson, S.S., and Zimmermann, R.A. (2002) Transit of tRNA through the *Escherichia coli* ribosome: cross-linking of the 3' end of tRNA to ribosomal proteins at the P and E sites. FEBS Lett. 514: 60-66.

Kleene, K.C., Cataldo, L., Mastrangelo, M.-A., and Tagne, J.-B. (2003) Alternative patterns of transcription and translation of the ribosomal protein L32 mRNA in somatic and spermatogenic cells in mice. Exp. Cell Res. 291: 101-110.

Klein, D.J., Moore, P.B., and Steitz, T.A. (2004) The roles of ribosomal proteins in the structure, assembly, and evolution of the large ribosomal subunit. J. Mol. Biol. 340: 141-177.

Klein, D.J., Schmeing, T.M., Moore, P.B., and Steitz, T.A. (2001) The kink-turn: a new RNA secondary structure motif. EMBO J. 20: 4214-4221.

Köhrer, C., Mayer, C., Neumair, O., Gröbner, P., and Piendl, W. (1998) Interaction of ribosomal L1 proteins from mesophilic and thermophilic Archaea and Bacteria with specific L1-binding sites on 23S rRNA and mRNA. Eur. J. Biochem. 256: 97-105.

Kongsuwan, K., Yu, Q., Vincent, A., Frisardi, M.C., Rosbash, M., Lengyel, J.A., and Merriam, J. (1985) A *Drosophila Minute* gene encodes a ribosomal protein. Nature 317: 555-558.

Kooi, E.A., Rutgers, C.A., Kleijmeer, M.J., van't Riet, J., Venema, J., and Raué, H.A. (1994) Mutational analysis of the C-terminal region of *Saccharomyces cerevisiae* ribosomal protein L25 *in vitro* and *in vivo* demonstrates the presence

of two distinct functional elements. J Mol. Biol. 240: 243-255.

Kooi, E.A., Rutgers, C.A., Mulder, A., van't Riet, J., Venema, J., and Raué, H.A. (1993) The phylogenetically conserved doublet tertiary interaction in domain III of the large subunit rRNA is crucial for ribosomal protein binding. Proc. Natl. Acad. Sci. U.S.A. 90: 213-216.

Köpke, A.K.E., Leggatt, P.A., and Matheson, A.T. (1992) Structure function relationships in the ribosomal stalk proteins of archaebacteria. J. Biol. Chem. 267: 1382-1390.

Kramer, G., Rauch, T., Rist, W., Vorderwülbecke, S., Patzelt, H., Schulze-Specking, A., Ban, N., Deuerling, E., and Bukau, B. (2002) L23 protein functions as a chaperone docking site on the ribosome. Nature 419: 171-174.

Lafontaine, D.L.J., and Tollervey, D. (1998) Birth of the snoRNPs: the evolution of the modification-guide snoRNAs. Trends Biochem. Sci. 23: 383-388.

Larkin, J.C., Hunsperger, J.P., Culley, D., Rubenstein, I., and Silflow, C.D. (1989) The organization and expression of a maize ribosomal protein gene family. Genes & Dev. 3: 500-509.

Larson, D.E., Zahradka, P., and Sells, B.H. (1991) Control points in eucaryotic ribosome biogenesis. Biochem. Cell Biol. 69: 5-22.

Lascaris, R.F., de Groot, E., Hoen, P.-B., Mager, W.H., and Planta, R.J. (2000) Different roles for Abf1p and a T-rich promoter element in nucleosome organization of the yeast *RPS28A* gene. Nuc. Acids Res. 28: 1390-1396.

Lascaris, R.F., Mager, W.F., and Planta, R.J. (1999) DNA-binding requirements of the yeast protein Rap1p as selected *in silico* from ribosomal protein gene promoter sequences. Bioinformatics 15: 267-277.

Le Hir, H., Nott, A., and Moore, M.J. (2003) How introns influence and enhance eukaryotic gene expression. Trends Biochem. Sci. 28: 215-220.

Lebrun, M., and Freyssinet, G. (1991) Nucleotide sequence and characterization of a maize cytoplasmic ribosomal protein S11 cDNA. Plant Mol. Biol. 17: 265-268.

Lecompte, O., Ripp, R., Thierry, J.-C., Moras, D., and Poch, O. (2002) Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. Nuc. Acids Res. 30: 5382-5390.

Ledda, M., Di Croce, M., Bedini, B., Wannenes, F., Corvaro, M., Pilo Boyl, P., Caldarola, S., Loreni, F., and Amaldi, F. (2005) Effect of 3' UTR length on the translational regulation of 5'-terminal oligopyrimidine mRNAs. Gene 344: 213-220.

Lee, C.-K., Shibata, Y., Rao, B., Strahl, B.D., and Lieb, J.D. (2004) Evidence for nucleosome depletion at active regulatory regions genome-wide. Nat. Genet. 36: 900-905.

Lee, H.-S., Mun, J.-H., and Kim, S.-G. (1999) Characterization of cDNAs encoding cytoplasmic ribosomal protein L15 and L27a in petunia (*Petunia hybrida*): primary structures and coordinate expression. Gene 226: 155-163.

Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I., Zeitlinger, J., Jennings, E.G., Murray, H.L., Gordon, D.B., Ren, B., Wyrick, J.J., Tagne J.-B., Volkert, T.L., Fraenkel, E., Gifford, D.K., and Young, R.A. (2002) Transcriptional regulatory networks in *Saccharomyces cerevisiae*. Science 298: 799-804.

Lenvik, T.R., Key, J.L., and Gantt, J.S. (1994) Ribosomal protein S11 genes from *Arabidopsis* and soybean. Plant Physiol. 105: 1027-1028.

Li, B., Nierras, C.R., and Warner, J.R. (1999) Transriptional elements involved in the repression of ribosomal protein synthesis. Mol. Cell. Biol. 19: 5393-5404.

Liao, S., Lin, J., Do, H., and Johnson, A.E. (1997) Both luminal and cytosolic gating of the aqueous ER translocon pore are regulated from inside the ribosome during membrane protein integration. Cell 90: 31-41.

Lieb, J.D., Liu, X., Botstein, D., and Brown, P.O. (2001) Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. Nat. Genet. 28: 327-334.

Lill, R., Crooke, E., Guthrie, B., and Wickner, W. (1988) The "trigger factor cycle" includes ribosomes, presecretory proteins, and the plasma membrane. Cell 54: 1013-1018.

Lin, C.T., and Thomashow, M.F. (1992) DNA sequence analysis of a complementary DNA for cold-regulated Arabidopsis gene *cor15* and characterization of the COR15 polypeptide. Plant Physiol. 99: 519-525.

Lindahl, L., and Zengel, J.M. (1986) Ribosomal genes in *Escherichia coli*. Ann. Rev. Genet. 20: 297-326.

Link, A.J., Eng, J., Schieltz, D.M., Carmack, E., Mize, G.J., Morris, D.R., Garvik, B.M., and Yates, J.R.III. (1999) Direct analysis of protein complexes using mass spectrometry. Nat. Biotech. 17: 676-682.

Littlefield, J.W., Keller, E.B., Gross, J., and Zamecnik, P.C. (1955) Studies on cytoplasmic ribonucleoprotein particles from the liver of the rat. J. Biol. Chem. 217: 111-123.

Long, E.O., and Dawid, I.B. (1980) Repeated genes in eukaryotes. Ann. Rev. Biochem. 49: 727-764.

Loreni, F., and Amaldi, F. (1992) Translational regulation of ribosomal protein synthesis in *Xenopus* cultured cells: mRNA relocation between polysomes and RNP during nutritional shifts. Eur. J. Biochem. 205: 1027-1032.

Ludwig, A., and Tenhaken, R. (2001) Suppression of the ribosomal *L2* gene reveals a novel mechanism for stress adaptation in soybean. Planta 212: 792-798.

Maden, B.E.H., and Monro, R.E. (1968) Ribosome-catalyzed peptidyl transfer: effects of cations and pH value. Eur. J. Biochem. 6: 309-316.

Maden, B.E.H., Traut, R.R., and Monro, R.E. (1968) Ribosome-catalyzed peptidyl transfer: the polyphenylalanine system. J. Mol. Biol. 35: 333-345.

Mager, W.H. (1988) Control of ribosomal protein gene expression. Biochim. et Biophys. Acta 949: 1-15.

Mager, W.H., Planta, R.J., Ballesta, J.-P.G., Lee, J.C., Mizuta, K., Suzuki, K., Warner, J.R., and Woolford, J. (1997) A new nomenclature for the cytoplasmic ribosomal proteins of *Saccharomyces cerevisiae*. Nuc. Acids Res. 25: 4872-4875.

Maier, T., Ferbitz, L., Deuerling, E., and Ban, N. (2005) A cradle for new proteins: trigger factor at the ribosome. Curr. Opin. Struct. Biol. 15: 204-212.

Maki, C.G., Rhoads, D.D., Diaz, J.-J., and Roufa, D.J. (1990) A *Drosophila* ribosomal protein functions in mammalian cells. Mol. Cell. Biol. 10: 4524-4528.

Mandiyan, V., Tumminia, S.J., Wall, J.S., Hainfield, J.F., and Boublik, M. (1991) Assembly of the *Escherichia coli* 30S ribosomal subunit reveals protein-dependent folding of the 16S rRNA domains. Proc. Natl. Acad. Sci. U.S.A. 88:

8174-8178.

Manevski, A., Bardet, C., Trémousaygue, D., and Lescure, B. (1999) Characterization and properties of heteromeric plant protein complexes that interact with *tef cis*-acting elements in both RNA polymerase II-dependent promoters and rDNA spacer sequences. Mol. Gen. Genet. 261: 892-900.

Manevski, A., Bertoni, G., Bardet, C., Trémousaygue, D., and Lescure, B. (2000) In synergy with various *cis*-acting elements, plant interstitial telomere motifs regulate gene expression in *Arabidopsis* root meristems. FEBS Lett. 483: 43-46.

Manuell, A.L., Yamaguchi, K., Haynes, P.A., Milligan, R.A., and Mayfield, S.P. (2005) Composition and structure of the 80S ribosome from the green alga *Chlamydomonas reinhardtii*: 80S ribosomes are conserved in plants and animals. J. Mol. Biol. 351: 266-279.

Mao, H., White, S.A., and Williamson, J.R. (1999) A novel loop-loop recognition motif in the yeast ribosomal protein L30 autoregulatory RNA complex. Nat. Struct. Biol. 6: 1139-1147.

Martin, D.E., and Hall, M.N. (2005) The expanding TOR signaling network. Curr. Opin. Cell Biol. 17: 158-166.

Martin, D.E., Soulard, A., and Hall, M.N. (2004) TOR regulates ribosomal protein gene expression via PKA and the forkhead transcription factor FHL1. Cell 119: 969-979.

Martin, T., Wöhner, R.-V., Hummel, S., Willmitzer, L., and Frommer, W.B. (1992) The GUS reporter system as a tool to study plant gene expression. In: Gallagher SR (Ed.) GUS protocols: using the GUS gene as a reporter of gene expression. Academic Press, San Diego, pp 23-43.

Marty, I., and Meyer, Y. (1992) cDNA nucleotide sequence and expression of a tobacco cytoplasmic ribosomal protein L2 gene. Nucl. Acids Res. 20: 1517-1522.

Matlack, K.E.S., and Walter, P. (1995) The 70 carboxyl-terminal amino acids of nascent secretory proteins are protected from proteolysis by the ribosome and the protein translocation apparatus of the endoplasmic reticulum membrane. J. Biol. Chem. 270: 6170-6180.

Matsumoto, K., Montzka Wassarman, K., and Wolffe, A.P. (1998) Nuclear history of a

pre-mRNA determines the translational activity of cytoplasmic mRNA. EMBO J. 17: 2107-2121.

Mazumder, B., Sampath, P., Seshadri, V., Maitra, R.K., DiCorleto, P.E., and Fox, P.L. (2003) Regulated release of L13a from the 60S ribosomal subunit as a mechanism of transcript-specific translational control. Cell 115: 187-198.

McBride, K.E., and Summerfelt, K.R. (1990) Improved binary vectors for *Agrobacterium*-mediated plant transformation. Plant Mol Biol 14:269-276.

McConkey, E.H., Bielka, H., Gordon, J., Lastick, S.M., Lin, A., Ogata, K., Reboud, J.-P., Traugh, J.A., Traut, R.R., Warner, J.R., Welfle, H., and Wool, I.G. (1979) Proposed uniform nomenclature for mammalian ribosomal proteins. Mol. Gen. Genet. 169: 1-6.

McIntosh, K.B, and Bonham-Smith, P.C. (2001) Establishment of *Arabidopsis thaliana* ribosomal protein RPL23AA as a functional homologue of *Saccharomyces cerevisiae* ribosomal protein L25. Plant Mol. Biol. 46: 673-682.

McIntosh, K.B., and Bonham-Smith, P.C. (2005) The two ribosomal protein *L23A* genes are differentially transcribed in *Arabidopsis thaliana*. Genome 48: 443-454.

Mears, J.A., Cannone, J.J., Stagg, S.M., Gutell, R.R., Agrawal, R.K., and Harvey, S.C. (2002) Modeling a minimal ribosome based on comparative sequence analysis. J. Mol. Biol.321: 215-234.

Melanson, D.L., and Ingle, J. (1978). Regulation of ribosomal RNA accumulation by auxin in artichoke tissue. Plant Physiol. 62: 761-765.

Menand, B., Desnos, T., Nussaume, L., Berger, F., Bouchez, D., Meyer, C., and Robaglia, C. (2002) Expression and disruption of the *Arabidopsis TOR* (target of rapamycin) gene. Proc. Natl. Acad. Sci. U.S.A. 99: 6422-6427.

Ménétret, J.-F., Neuhof, A., Morgan, D.G., Plath, K., Radermacher, M., Rapoport, T., and Akey, C.W. (2000) The structure of ribosome-channel complexes engaged in protein translocation. Mol. Cell 6: 1219-1232.

Merianos, H.J., Wang, J., and Moore, P.B. (2004) The structure of a ribosomal protein S8/*spc* operon mRNA complex. RNA 10: 954-964.

Metzenberg, S., Joblet, C., Verspieren, P., and Agabian, N. (1993) Ribosomal protein L25 from *Trypanosoma brucei*: phylogeny and molecular co-evolution of an

rRNA-binding protein and its rRNA-binding site. Nuc. Acids Res. 21: 4936-4940.

Meyuhas, O. (2000) Synthesis of the translational apparatus is regulated at the translational level. Eur. J. Biochem. 267: 6321-6330.

Meyuhas, O., and Klein, A. (1990) The mouse ribosomal protein L7 gene. J. Biol. Chem. 265: 11465-11473.

Miller, O.L. (1981) The nucleolus, chromosomes, and visualization of genetic activity. J. Cell Biol. 91: 15s-27s.

Mitrovich, Q.M., and Anderson, P. (2000) Unproductively spliced ribosomal protein mRNAs are natural targets of mRNA surveillance in *C. elegans*. Genes Dev. 14: 2173-2184.

Miyake, K., Flygare, J., Kiefer, T., Utsugisawa, T., Richter, J., Ma, Z., Wiznerowicz, M., Trono, D., and Karlsson, S. (2005) Development of cellular models for ribosomal protein S19 (RPS19)-deficient Diamond-Blackfan anemia using inducible expression of siRNA against RPS19. Mol. Therapy 11: 627-637.

Mizoguchi, T., Hayashida, N., Yamaguchi-Shinozaki, K., Kamada, H., and Shinozaki, K. (1995) Two genes that encode ribosomal-protein S6 kinase homologs are induced by cold or salinity stress in *Arabidopsis thaliana*. FEBS Lett. 358: 199-204.

Mizuta, K., and Warner, J.R. (1994) Continued functioning of the secretory pathway is essential for ribosome synthesis. Mol. Cell. Biol. 14: 2493-2502.

Monro, R.E. (1967) Catalysis of peptide bond formation by 50S ribosomal subunits from *Escherichia coli*. J. Mol. Biol. 26: 147-151.

Moore, P.B. (1988) The ribosome returns. Nature 331: 223-227.

Moore, P.B. (1993) Ribosomes and the RNA world. In: The RNA World. Edited by: Gesteland, R.F., and Atkins, J.F. Cold Spring Harbor Laboratory Press, New York. pp.119-135.

Moore, P.B., and Steitz, T.A. (2002) The involvement of RNA in ribosome function. Nature 418: 229-235.

Moore, V.G., Atchison, R.E., Thomas, G., Moran, M., and Noller, H.F. (1975) Identification of a ribosomal protein essential for peptidyl transferase activity.

Proc. Natl. Acad. Sci. U.S.A. 72: 844-848.

Moran, D.L. (2000) Characterization of the structure and expression of a highly conserved ribosomal protein gene, L9, from pea. Gene 253: 19-29.

Morgan, D.G., Ménétret, J.-F., Neuhof, A., Rapoport, T.A., and Akey, C.W. (2002) Structure of the mammalian ribosome-channel complex at 17 Å resolution. J. Mol. Biol. 324: 871-886.

Morgan, D.G., Ménétret, J.-F., Radermacher, M., Neuhof, A., Akey, I.V., Rapoport, T., and Akey, C.W. (2000) A comparison of the yeast and rabbit 80S ribosome reveals the topology of the nascent chain exit tunnel, inter-subunit bridges and mammalian rRNA expansion segments. J. Mol. Biol. 301: 301-321.

Murashige, T., and Skoog, F. (1962) A revised medium for rapid growth and bio-assays with tobacco tissue cultures. Physiol. Planta. 15: 473-497.

Murphy, F.V., and Ramakrishnan, V. (2004) Structure of a purine-purine wobble base pair in the decoding center of the ribosome. Nat. Struct. Mol. Biol. 11: 1251-1252.

Nakao, A., Yoshihama, M., and Kenmochi, N. (2004) RPG: the Ribosomal Protein Gene database. Nuc. Acids Res. 32: D168-D170.

Nakatogawa, H., and Ito, K. (2002) The ribosomal exit tunnel functions as a discriminating gate. Cell 108: 629-636.

Naora, H., and Naora, H. (1999) Involvement of ribosomal proteins in regulating cell growth and apoptosis: translational modulation or recruitment for extraribosomal activity? Immunol. Cell Biol. 77: 197-205.

Nazar, R.N. (1980) A 5.8S rRNA-like sequence in prokaryotic 23S rRNA. FEBS Lett. 119: 212-214.

Nevskaya, N., Tishchenko, S., Gabdoulkhakov, A., Nikonova, E., Nikonov, O., Nikulin, A., Platonova, O., Garber, M., Nikonov, S., and Piendl, W. (2005) Ribosomal protein L1 recognizes the same specific structural motif in its target sites on the autoregulatory mRNA and 23S rRNA. Nuc. Acids Res. 33: 478-485.

Nierhaus, K.H., and Montejo, V. (1973) A protein involved in the peptidyl transferase activity of *Escherichia coli* ribosomes. Proc. Natl. Acad. Sci. U.S.A. 70: 1931-1935.

Nierhaus, K.H. (1991) The assembly of prokaryotic ribosomes. Biochimie 73: 739-755.

Nikulin, A., Eliseikina, I., Tishchenko, S., Nevskaya, N., Davydova, N., Platonova, O., Piendl, W., Selmer, M., Liljas, A., Drygin, D., Zimmermann, R., Garber, M., and Nikonov, S. (2003) Structure of the L1 protuberance in the ribosome. Nat. Struct. Biol. 10: 104-108.

Nilsson, J., Sengupta, J., Frank, J., and Nissen, P. (2004) Regulation of eukaryotic translation by the RACK1 protein: a platform for signalling molecules on the ribosome. EMBO Rep. 5: 1137-1141.

Nirenberg, M.W., and Matthaie, J.H. (1961) The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. Proc. Natl. Acad. Sci. U.S.A. 47: 1588-1602.

Nirenberg, M., Leder, P., Bernfield, M., Brimacombe, R., Trupin, J., Rottman, F., and O'Neal, C. (1965) RNA codewords and protein synthesis, VII. On the general nature of the RNA code. Proc. Natl. Acad. Sci. U.S.A. 53: 1161-1168.

Nishimura, T., Wada, T., and Okada, K. (2004) A key factor of translation reinitiation, ribosomal protein L24, is involved in gynoecium development in *Arabidopsis*. Biochem. Soc. Trans. 32: 611-613.

Nissan, T.A., Baßler, J., Petfalski, E., Tollervey, D., and Hurt, E. (2002) 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. EMBO J. 21: 5539-5547.

Nissen, P., Hansen, J., Ban, N., Moore, P.B., and Steitz, T.A. (2000) The structural basis of ribosome activity in peptide bond synthesis. Science 289: 920-930.

Noller, H.F. (1991) Ribosomal RNA and translation. Ann. Rev. Biochem. 60: 191-227.

Noller, H.F., Hoffarth, V., and Zimniak, L. (1992) Unusual resistance of peptidyl transferase to protein extraction procedures. Science 256: 1416-1419.

Noller, H.F. (1993) Peptidyl transferase: protein, ribonucleoprotein, or RNA? J. Bacteriol. 175: 5297-5300.

Nomura, M. (1990) History of ribosome research: a personal account. In: The Ribosome: Structure, Function, and Evolution. Edited by: Hill W.E., Dahlberg, A., Garrett, R.A.,Moore, P.B., Schlessinger, D., and Warner, J.R. American Society for Microbiology, Washington, D.C. pp.3-55.

Nomura, M. (1999) Regulation of ribosome biosynthesis in *Escherichia coli* and *Saccharomyces cerevisiae*: diversity and common principles. J. Bacteriol. 181: 6857-6864.

Nomura, M., Yates, J.L., Dean, D., and Post, L.E. (1980) Feedback regulation of ribosomal protein gene expression in *Escherichia coli*: structural homology of ribosomal RNA and ribosomal protein mRNA. Proc. Natl. Acad. Sci. U.S.A. 77: 7084-7088.

Nott, A., Le Hir, H., and Moore, M.J. (2004) Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. Genes & Dev. 18: 210-222.

Nover, L., Munsche, D., Ohme, K., and Scharf, K.-D. (1986) Control of ribosome biosynthesis in plant cell cultures under heat-shock conditions. Eur. J. Biochem. 160: 297-304.

Nowotny, V., and Nierhaus, K.H. (1982) Initiator proteins for the assembly of the 50S subunit from *Escherichia coli* ribosomes. Proc. Natl. Acad. Sci. U.S.A. 79: 7238-7242.

Nusspaumer, G., Remacha, M., and Ballesta, J.P. (2000) Phosphorylation and N-terminal region of yeast ribosomal protein P1 mediate its degradation, which is prevented by protein P2. EMBO J. 19: 6075-6084.

Odintsova, T.I., Müller, E.-C., Ivanov, A.V., Egorov, T.A., Bienert, R., Vladimirov, S.N., Kostka, S., Otto, A., Wittmann-Liebold, B., and Karpova, G.G. (2003) Characterization and analysis of posttranslational modifications of the human large cytoplasmic ribosomal subunit proteins by mass spectrometry and Edman sequencing. J. Prot. Chem. 22: 249-258.

Ogle, J.M., Brodersen, D.E., Clemons, W.M. Jr., Tarry, M.J., Carter, A.P., and Ramakrishnan, V. (2001) Recognition of cognate transfer RNA by the 30S ribosomal subunit. Science 292: 897-902.

Ohta, S., Mitra, S., Hattori, T., and Nakamura, K. (1990) Construction and expression in tobacco of a β-glucuronidase (GUS) reporter gene containing an intron within the coding sequence. Plant Cell Physiol. 31:805-813.

Oubridge, C., Ito, N., Evans, P.R., Teo, C.-H., and Nagal, K. (1994) Crystal structure at

1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. Nature 372: 432-438.

Palade, G.E. (1955) A small particulate component of the cytoplasm. J. Biophys. Biochem. Cytol. 1: 59-68.

Palade, G.E., and Siekevitz, P. (1956) Liver microsomes: an integrated morphological and biochemical study. J. Biophys. Biochem. Cytol. 2: 171-200.

Panopoulos, P., Dresios, J., and Synetos, D. (2004) Biochemical evidence of translational infidelity and decreased peptidyltransferase activity by a sarcin/ricin domain mutation of yeast 25S rRNA. Nuc. Acids Res. 32: 5398-5408.

Patterson, A., Karsi, A., Feng, J., and Liu, Z. (2003) Translational machinery of channel catfish: II. Complementary DNA and expression of the complete set of 47 60S ribosomal proteins. Gene 305: 151-160.

Pennisi, E. (1999) The race to the ribosome structure. Science 285: 2048-2051.

Pérez, L., Aguilar, R., Pérez-Méndez, A., and Sánchez de Jiménez, E. (1990) Phosphorylation of ribosomal proteins induced by auxins in maize embryonic tissues. Plant Physiol. 94: 1270-1275.

Pérez-Méndez, A., Aguilar, R., Briones, E., and Sánchez de Jiménez, E. (1993) Characterization of ribosomal protein phosphorylation in maize axes during germination. Plant Sci. 94: 71-79.

Perry, R.P. (2005) The architecture of mammalian ribosomal protein promoters. BMC Evol. Biol. 5:15.

Petermann, M.L., and Hamilton, M.G. (1957) The purification and properties of cytoplasmic ribonucleoprotein from rat liver. J. Biol. Chem. 224: 725-736.

Pierandrei-Amaldi, P., Beccari, E., Bozzoni, I., and Amaldi, F. (1985) Ribosomal protein production in normal and anucleolate *Xenopus* embryos: regulation at the posttranscriptional and translational levels. Cell 42: 317-323.

Pierandrei-Amaldi, P., Campioni, N., Beccari, E., Bozzoni, I., and Amaldi, F. (1982) Expression of ribosomal-protein genes in Xenopus laevis development. Cell 30: 163-171.

Planta, R.J., Gonçalves, P.M., and Mager, W.H. (1995) Global regulators of ribosome biosynthesis in yeast. Biochem. Cell Biol. 73: 825-834.

Planta, R.J., and Mager, W.M. (1998) The list of cytoplasmic ribosomal proteins of *Saccharomyces cerevisiae*. Yeast 14: 471-477.

Polacek, N., Gaynor, M., Yassin, A., and Mankin, A.S. (2001) Ribosomal peptidyl transferase can withstand mutations at the putative catalytic nucleotide. Nature 411: 498-501.

Pool, M.R., Stumm, J., Fulga, T.A., Sinning, I., and Dobberstein, B. (2002) Distinct modes of signal recognition particle interaction with the ribosome. Science 297: 1345-1348.

Poole, A., Jeffares, D., and Penny, D. (1999) Early evolution: prokaryotes, the new kids on the block. BioEssays 21: 880-889.

Popescu, S.C., and Tumer, N.E. (2004) Silencing of ribosomal protein L3 genes in *N. tabacum* reveals coordinate expression and significant alterations in plant growth, development and ribosome biogenesis. Plant J. 39: 29-44.

Potter, M.D., and Nicchitta, C.V. (2000) Regulation of ribosome detachment from the mammalian endoplasmic reticulum membrane. J. Biol. Chem. 275: 33828-33835.

Potter, M.D., and Nicchitta, C.V. (2002) Endoplasmic reticulum-bound ribosomes reside in stable association with the translocon following termination of protein synthesis. J. Biol. Chem. 277: 23314-23320.

Powers, T., and Walter, P. (1999) Regulation of ribosome biogenesis by the rapamycin-sensitive TOR-signaling pathway in *Saccharomyces cerevisiae*. Mol. Biol. Cell 10: 987-1000.

Prestridge, D.S. (1991) SIGNAL SCAN: a computer program that scans DNA sequences for eukaryotic transcriptional elements. CABIOS 7: 203-206.

Presutti, C., Villa, T., Hall, D., Pertica, C., and Bozzoni, I. (1995) Identification of the *cis*-elements mediating the autogenous control of ribosomal protein L2 mRNA stability in yeast. EMBO J. 14: 4022-4030.

Prinz, A., Behrens, C., Rapoport, T.A., Hartmann, E., and Kalies, K.-U. (2000) Evolutionarily conserved binding of ribosomes to the translocation channel via the large ribosomal RNA. EMBO J. 19: 1900-1906.

Raibaud, S., Vachette, P., Guillier, M., Allemand, F., Chiaruttini, C., and Dardel, F.

(2003) How bacterial ribosomal protein L20 assembles with 23S ribosomal RNA and its own messenger RNA. J. Biol. Chem. 278: 36522-36530.

Rakouský, S., Kocábek, T., Vincenciová, R., and Ondrej, M. (1997) Transient β-glucuronidase activity after infiltration of *Arabidopsis thaliana* by *Agrobacterium tumefaciens*. Biol. Plant. 40:33-41.

Ramakrishnan, V., and Moore, P.B. (2001) Atomic structures at last: the ribosome in 2000. Curr. Opin. Struct. Biol. 11: 144-154.

Ramakrishnan, V., Davies, C., Gerchman, S.E., Golden, B.L., Hoffmann, D.W., Jaishree, T.N., Kycia, J.H., Porter, S., and White, S.W. (1995) Structures of prokaryotic ribosomal proteins: implications for RNA binding and evolution. Biochem. Cell Biol. 73: 979-986.

Ramakrishnan, V., and White, S.W. (1992) The structure of ribosomal protein S5 reveals sites of interaction with 16S rRNA. Nature 358: 768-771.

Raué, H.A., Klootwijk, J., and Musters, W. (1988) Evolutionary conservation of structure and function of high molecular weight ribosomal RNA. Prog. Biophys. Mol. Biol. 51: 77-129.

Regad, F., Hervé, C., Marinx, O., Bergounioux, C., Trémousaygue, D., and Lescure, B. (1995) The *tef1* box, a ubiquitous *cis*-acting element involved in the activation of plant genes that are highly expressed in cycling cells. Mol. Gen. Genet. 248: 703-711.

Reid, J.L., Iyer, V.R., Brown, P.O., and Struhl, K. (2000) Coordinate regulation of yeast ribosomal protein genes is associated with targeted recruitment of Esa1 histone acetylase. Mol. Cell 6: 1297-1307.

Reiter, A.K., Anthony, T.G., Anthony, J.C., Jefferson, L.S., and Kimball, S.R. (2004) The mTOR signaling pathway mediates control of ribosomal protein mRNA translation in rat liver. Int. J. Biochem. Cell Biol. 36: 2169-2179.

Remacha, M., Jimenez-Diaz, A., Bermejo, B., Rodríguez-Gabriel, M.A., Guarinos, E., and Ballesta, J.P.G. (1995) Ribosomal acidic phosphoproteins P1 and P2 are not required for cell viability but regulate the pattern of protein expression in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 15: 4754-4762.

Revenkova, E., Masson, J., Koncz, C., Afsar, K., Jakovleva, L., and Paszkowski, J.

(1999) Involvement of *Arabidopsis thaliana* ribosomal protein S27 in mRNA degradation triggered by genotoxic stress. EMBO J. 18: 490-499.

Rhoads, D.D., and Roufa, D.J. (1987) A cloned human ribosomal protein gene functions in rodent cells. Mol. Cell. Biol. 7: 3767-3774.

Richards, E.J., and Ausubel, F.M. (1988) Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. Cell 53: 127-136.

Robert, F., and Brakier-Gingras, L. (2003) A functional interaction between ribosomal proteins S7 and S11 within the bacterial ribosome. J. Biol. Chem. 278: 44913-44920.

Roberts, R.B. (1958) Microsomal Particles and Protein Synthesis. New York: Pergamon Press.

Rodnina, M.V., and Wintermeyer, W. (2003) Peptide bond formation on the ribosome: structure and mechanism. Curr. Opin. Struct. Biol. 13: 334-340.

Rodríguez-Gabriel, M.A., Remacha, M., and Ballesta, J.P.G. (2000) The RNA interacting domain but not the protein interacting domain is highly conserved in ribosomal protein P0. J. Biol. Chem. 275: 2130-2136.

Rogers, H.J., Bate, N., Combo, J., Sullivan, J., Sweetman, J., Swan, C., Lonsdale, D.M., and Twell, D. (2001) Functional analysis of *cis*-regulatory elements within the promoter of the tobacco late pollen gene g10. Plant Mol. Biol. 45: 577-585.

Rohde, J.R., and Cardenas, M.E. (2003) The Tor pathway regulates gene expression by linking nutrient sensing to histone acetylation. Mol. Cell. Biol. 23: 629-635.

Röhl, R., and Nierhaus, K.H. (1982) Assembly map of the large subunit (50S) of *Escherichia coli* ribosomes. Proc. Natl. Acad. Sci. U.S.A. 79: 729-733.

Rose, A.B. (2004) The effect of intron location on intron-mediated enhancement of gene expression in *Arabidopsis*. Plant J. 40: 744-751.

Rossi, L., Escudero, J., Hohn, B., and Tinland, B. (1993) Efficient and sensitive assay for T-DNA-dependent transient gene expression. Plant Mol. Biol. Rep. 11:220-229.

Rotenberg, M.O., and Woolford, J.L.Jr. (1986) Tripartite upstream promoter element essential for expression of *Saccharomyces cerevisiae* ribosomal protein genes. Mol. Cell. Biol. 6: 674-687.

Rudra, D., Zhao, Y., and Warner, J.R. (2005) Central role of Ifh1p-Fhl1p interaction in the synthesis of yeast ribosomal proteins. EMBO J. 24: 533-542.

Rutgers, C.A., Schaap, P.J., van't Riet, J., Woldringh, C.L., and Raué, H.A. (1990) *In vivo* and *in vitro* analysis of structure-function relationships in ribosomal protein L25 from *Saccharomyces cerevisiae*. Biochim. et Biophys. Acta 1050: 74-79.

Rutgers, C.A., Rientjes, J.M.J., van't Riet, J., and Raué, H.A. (1991) rRNA binding domain of yeast ribosomal protein L25: identification of its borders and a key leucine residue. J. Mol. Biol. 218: 375-385.

Sáez-Vásquez, J., Gallois, P., and Delseny, M. (2000) Accumulation and nuclear targeting of BnC24, a *Brassica napus* ribosomal protein corresponding to a mRNA accumulating in response to cold treatment. Plant Sci. 156: 35-46.

Sanger, F., Miklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitor. Proc. Natl. Acad. Sci. U.S.A. 74: 5463-5467.

Sangwan, V., Lenvik, T.R., and Gantt, J.S. (1993) The *Arabidopsis thaliana* ribosomal protein S15 (*rig*) gene. Biochim. et Biophys. Acta 1216: 221-226.

Santos, C., Rodríguez-Gabriel, M.A., Remacha, M., and Ballesta, J.P.G. (2004) Ribosomal P0 protein domain involved in selectivity of antifungal sordarin derivatives. Antimicrob. Agents Chemother. 48: 2930-2936.

Saveanu, C., Namane, A., Gleizes, P.-E., Lebreton, A., Rousselle, J.-C., Noaillac-Depeyre, J., Gas, N., Jacquier, A., and Fromont-Racine, M. (2003) Sequential protein association with nascent 60S ribosomal particles. Mol. Cell. Biol. 23: 4449-4460.

Schaap, P.J., van't Riet, J., Woldringh, C.L., and Raué, H.A. (1991) Identification and functional analysis of the nuclear localization signals of ribosomal protein L25 from *Saccharomyces cerevisiae*. J. Mol. Biol. 221: 225-237.

Schäfer, T., Strauß, D., Petfalski, E., Tollervey, D., and Hurt, E. (2003) The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes. EMBO J. 22: 1370-1380.

Scharf, K.-D., and Nover, L. (1987) Control of ribosome biosynthesis in plant cell cultures under heat shock conditions II. Ribosomal proteins. Biochim. et Biophys. Acta 909: 44-57.

Scharf, K.-D., and Nover, L. (1982) Heat-shock-induced alterations of ribosomal

protein phosphorylation in plant cell cultures. Cell 30: 427-437.

Schawalder, S.B., Kabanl, M., Howald, I., Choudhury, U., Werner, M., and Shore, D. (2004) Growth-regulated recruitment of the essential yeast ribosomal protein gene activator Ifh1. Nature 432: 1058-1061.

Schirmer, E.C., Lindquist, S., and Vierling, E. (1994) An Arabidopsis heat shock protein complements a thermotolerance defect in yeast. Plant Cell 6:1899-1909.

Schnare, M.N., Damberger, S.H., Gray, M.W., and Gutell, R.R. (1996) Comprehensive comparison of structural characteristics in eukaryotic cytoplasmic large subunit (23S-like) ribosomal RNA. J. Mol. Biol. 256: 701-719.

Schöffl, F., Prändl, R., Reindl, A. (1998) Regulation of the heat-shock response. Plant Physiol. 117:1135-1141.

Schweizer, P., Pokorny, J., Abderhalden, O., and Dudler, R. (1999) A transient assay system for the functional assessment of defense-related genes in wheat. Mol. Plant Micr. Int. 12:647-654.

Schwindinger, W.F., and Warner, J.R. (1987) Transcriptional elements of the yeast ribosomal protein gene *CYH2*. J. Biol. Chem. 262: 5690-5695.

Sengupta, J., Nilsson, J., Gursky, R., Spahn, C.M.T., Nissen, P., and Frank, J. (2004) Identification of the versatile scaffold protein RACK1 on the eukaryotic ribosome by cryo-EM. Nat. Struct. Mol. Biol. 11: 957-962.

Sessa, G., Borello, U., Morelli, G., and Ruberti, I. (1998) A transient assay for rapid functional analysis of transcription factors in Arabidopsis. Plant Mol. Biol. Rep. 16:191-197.

Shama, S., and Meyuhas, O. (1996) The translational *cis*-regulatory element of mammalian ribosomal protein mRNAs is recognized by the plant translational apparatus. Eur. J. Biochm. 236: 383-388.

Sheen, J. (2001) Signal transduction in maize and Arabidopsis mesophyll protoplasts. Plant Physiol. 127:1466-1475.

Shen, X., Valencia, C.A., Szostak, J., Dong, B., and Liu, R. (2005) Scanning the human proteome for calmodulin-binding proteins. Proc. Natl. Acad. Sci. U.S.A. 102: 5969-5974.

Sherman, F., Fink, G.R., and Hicks, J.B. (1983) Methods in Yeast Genetics. Cold Spring

Harbor Laboratory Press, Cold Spring Harbor, NY.

Sherton, C.C., and Wool, I.G. (1972) Determination of the number of proteins in liver ribosomes and ribosomal subunits by two-dimensional polyacrylamide gel electrophoresis. J. Biol. Chem. 247: 4460-4467.

Shine, J., and Dalgarno, L. (1974) The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. U.S.A. 71: 1342-1346.

Shor, B., Calaycay, J., Rushbrook, J., and McLeod, M. (2003) Cpc2/RACK1 is a ribosome-associated protein that promotes efficient translation in *Schizosaccharomyces pombe*. J. Biol. Chem. 278: 49119-49128.

Sieburth, L.E., and Meyerowitz, E.M. (1997) Molecular dissection of the *AGAMOUS* control region shows that *cis* elements for spatial regulation are located intragenically. Plant Cell 9: 355-365.

Sievers, A., Beringer, M., Rodnina, M.V., and Wolfenden, R. (2004) The ribosome as an entropy trap. Proc. Natl. Acad. Sci. U.S.A. 101: 7897-7901.

Sikorski, M.M., Przybyl, D., Legocki, A.B., and Nierhaus, K.H. (1983) Group fractionation of wheat germ ribosomal proteins. Plant Sci. Lett. 30: 303-320.

Skriver, K., Olsen, F.L., Rogers, J.C., and Mundy, J. (1991) cis-acting DNA elements responsive to gibberellin and its antagonist abscisic acid. Proc. Natl. Acad. Sci. U. S.A. 88: 7266-7270.

Söll, D., Ohtsuka, E., Jones, D.S., Lohrmann, R., Hayatsu, H., Nishimura, S., and Khorana, H.G. (1965) Studies on polynucleotides, XLIX. Stimulation of the binding of aminoacyl-sRNA's to ribosomes by ribotrinucleotides and a survey of codon assignments for 20 amino acids. Proc. Natl. Acad. Sci. U.S.A. 54: 1378-1385.

Sonnemann, J., Bäuerle, A., Winckler, T., and Mutzel, R. (1991) A ribosomal calmodulin-binding protein from *Dictyostelium*. J. Biol. Chem. 266: 23091-23096.

Sollner-Webb, B., and Tower, J. (1986) Transcription of cloned eukaryotic ribosomal RNA genes. Ann. Rev. Biochem. 55: 801-830.

Spahn, C.M.T., Beckmann, R., Eswar, N., Penczek, P.A., Sali, A., Blobel, G., and Frank,

J. (2001) Structure of the 80S ribosome from *Saccharomyces cerevisiae* – tRNA-ribosome and subunit-subunit interactions. Cell 107: 373-386.

Spingola, M., Grate, L., Haussler, D., and Ares, M.Jr. (1999) Genome-wide bioinformatic and molecular analysis of introns in *Saccharomyces cerevisiae*. RNA 5: 221-234.

Spirin, A.S., and Gavrilova, L.P. (1969) The ribosome. Springer-Verlag, New York.

Srivastava, A.K., and Schlessinger, D. (1990) rRNA processing in *Escherichia coli*. In: The Ribosome: Structure, Function, and Evolution. Edited by: Hill W.E., Dahlberg, A., Garrett, R.A.,Moore, P.B., Schlessinger, D., and Warner, J.R. American Society for Microbiology, Washington, D.C. pp.426-434.

Srivastava, S., Verschoor, A., Radermacher, M., Grassucci, R., and Frank, J. (1995) Three-dimensional reconstruction of mammalian 40S ribosomal subunit embedded in ice. J. Mol. Biol. 245: 461-466.

Stagg, S.M., Mears, J.A., and Harvey, S.C. (2003) A structural model for the assembly of the 30S subunit of the ribosome. J. Mol. Biol. 328: 49-61.

Steitz, J.A., and Jakes, K. (1975) How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. 72: 4734-4738.

Steitz, T.A. (2005) On the structural basis of peptide-bond formation and antibiotic resistance from atomic structures of the large ribosomal subunit. FEBS Lett. 579: 955-958.

Steitz, T.A., and Moore, P.B. (2003) RNA, the first macromolecular catalyst: the ribosome is a ribozyme. Trends Biochem. Sci. 28: 411-418.

Stelzl, U., Zengel, J.M., Tovbina, M., Walker, M., Nierhaus, K.H., Lindahl, L., and Patel, D.J. (2003) RNA-structural mimicry in *Escherichia coli* ribosomal protein L4-dependent regulation of the *S10* operon. J. Biol. Chem. 278: 28237-28245.

Stewart, M.J., and Thomas, G. (1994) Mitogenesis and protein synthesis: a role for ribosomal protein S6 phosphorylation? BioEssays 16: 809-815.

Still, D.W., Dahal, P., and Bradford, K.J. (1997) A single-seed assay for endo-β-mannanase activity from tomato endosperm and radicle tissues. Plant Physiol.

113: 13-20.

Strafstrom, J.P., and Sussex, I.M. (1992) Expression of a ribosomal protein gene in the axillary buds of pea seedlings. Plant Physiol. 100: 1494-1502.

Stroud, R.M., and Walter, P. (1999) Signal sequence recognition and protein targeting. Curr. Opin. Struct. Biol. 9: 754-759.

Subramanian, A.R. (1975) Copies of proteins L7 and L12 and heterogeneity of the large subunit of *Escherichia coli* ribosome. J. Mol. Biol. 95: 1-8.

Subramanian, A.R., Smooker, P.M., and Giese, K. (1990) Chloroplast ribosomal proteins and their genes. In: The Ribosome: Structure, Function, and Evolution. Edited by: Hill, W.E., Dahlberg, A., Garrett, R.A., Moore, P.B., Schlessinger, D., and Warner, J.R. American Society for Microbiology, Washington, D.C. pp: 655-663.

Sung, D.Y., Vierling, E., Guy, C.L. (2001) Comprehensive expression profile analysis of the Arabidopsis Hsp70 gene family. Plant Physiol. 126: 789-800.

Sutoh, K., and Yamauchi, D. (2003) Two *cis*-acting elements necessary and sufficient for gibberellin-upregulated proteinase expression in rice seeds. Plant J. 34: 635-645.

Suzuki, K., and Wool, I.G. (1993) The primary structure of rat ribosomal protein L23a. J. Biol. Chem. 268: 2755-2761.

Szick, K., Springer, M., and Bailey-Serres, J. (1998) Evolutionary analyses of the 12-kDa acidic ribosomal P-proteins reveal a distinct protein of higher plant ribosomes. Proc. Natl. Acad. Sci. U.S.A. 95: 2378-2383.

Szick-Miranda, K., and Bailey-Serres, J. (2001) Regulated heterogeneity in 12-kDa P-protein phosphorylation and composition of ribosomes in maize (*Zea mays* L.). J. Biol. Chem. 276: 10921-10928.

Tal, M., Weissman, I., and Silberstein, A. (1990) A new method for stoichiometric analysis of proteins in complex mixture – reevaluation of the stoichiometry of *E. coli* ribosomal proteins. J. Biochem. Biophys. Methods 21: 247-266.

Tasheva, E.S., and Roufa, D.J. (1995) Regulation of human *RPS14* transcription by intronic antisense RNAs and ribosomal protein S14. Genes & Dev. 9: 304-316.

Tatematsu, K., Ward, S., Leyser, O., Kamiya, Y., and Nambara, E. (2005) Identification

of cis-elements that regulate gene expression during initiation of axillary bud outgrowth in Arabidopsis. Plant Physiol. 138: 757-766.

Taylor, M.A., Mad Arif, S.A., Pearce, S.R., Davies, H.V., Kumar, A., and George, L.A. (1992) Differential expression and sequence analysis of ribosomal protein genes induced in stolon tips of potato (*Solanum tuberosum* L.) during the early stages of tuberisation. Plant Physiol. 100: 1171-1176.

Teraoka, H., and Nierhaus, K.H. (1978) Protein L16 induces a conformational change when incorporated into an L16-deficient core derived from *Escherichia coli* ribosomes. FEBS Lett. 88: 223-226.

Thiede, B., Urlauβ, H., Neubauer, H., Grelle, G., and Wittmann-Liebold, B. (1998) Precise determination of RNA-protein contact sites in the 50S ribosomal subunit of *Escherichia coli*. Biochem. J. 334: 39-42.

Thomas, M.J., and Seto, E. (1999) Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key? Gene 236: 197-208.

Thompson, J.D., Higgins, D.G., Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nuc. Acids Res. 22: 4673-4680.

Tishchenko, S., Nikulin, A., Fomenkova, N., Nevskaya, N., Nikonov, O., Dumas, P., Moine, H., Ehresmann, B., Ehresmann, C., Piendl, W., Lamzin, V., Garber, M., and Nikonov, S. (2001) Detailed analysis of RNA-protein interactions within the ribosoml protein S8-rRNA complex from the archaeon *Methanococcus jannaschii*. J. Mol. Biol. 311: 311-324.

Toorop, P.E., Barroco, R.M., Engler, G., Groot, S.P.C., and Hilhorst, H.W.M. (2005) Differentially expressed genes associated with dormancy or germination of *Arabidopsis thaliana* seeds. Planta 221: 637-647.

Torres, M., Condon, C., Balada, J.-M., Squires, C., and Squires, C.L. (2001) Ribosomal protein S4 is a transcription factor with properties remarkably similar to NusA, a protein involved in both non-ribosomal and ribosomal RNA antitermination. EMBO J. 20: 3811-3820.

Trémousaygue, D., Garnier, L., Bardet, C., Dabos, P., Hervé, C., and Lescure, B. (2003)

Internal telomeric repeats and 'TCP domain' protein-binding sites co-operate to regulate gene expression in *Arabidopsis thaliana* cycling cells. Plant J. 33: 957-966.

Trémousaygue, D., Manevski, A., Bardet, C., Lescure, N., and Lescure, B. (1999) Plant interstitial telomere motifs participate in the control of gene expression in root meristems. Plant J. 20: 553-561.

Trotta, C.R., Lund, E., Kahan, L., Johnson, A.W., and Dahlberg, J.E. (2003) Coordinated nuclear export of 60S ribosomal subunits and NMD3 in vertebrates. EMBO J. 22: 2841-2851.

Ts'o, P.O.P., Bonner, J., and Vinograd, J. (1956) Microsomal nucleoprotein particles from pea seedlings. J. Biophys. Biochem. Cytol. 2: 451-466.

Ts'o, P.O.P. (1962) The ribosomes – ribonucleoprotein particles. Ann. Rev. Plant Physiol. 13: 45-80.

Tsurugi, K., Collatz, E., Todokoro, T., and Wool, I.G. (1977) Isolation of eukaryotic ribosomal proteins. Purification and characterization of 60S ribosomal subunit proteins L3, L6, L7', L8, L10, L15, L17, L18, L19, L23', L25, L27', L28, L29, L31, L32, L34, L35, L36, L36', and L37'. J. Biol. Chem. 252: 3961-3969.

Turck, F., Kozma, S.C., Thomas, G., and Nagy, F. (1998) A heat-sensitive *Arabidopsis thaliana* kinase substitutes for human $p70^{s6k}$ function in vivo. Mol. Cell. Biol. 18: 2038-2044.

Turck, F., Zilbermann, F., Kozma, S.C., Thomas, G., and Nagy, F. (2004) Phytohormones participate in an S6 kinase signal transduction pathway in Arabidopsis. Plant Physiol. 134: 1527-1535.

Uchiumi, T., Hori, K., Nomura, T., and Hachimori, A. (1999) Replacement of L7/L12.L10 protein complex in *Escherichia coli* ribosomes with the eukaryotic counterpart changes the specificity of elongation factor binding. J. Biol. Chem. 274: 27578-27582.

Uechi, T., Tanaka, T., and Kenmochi, N. (2001) A complete map of the human ribosomal protein genes: assignment of 80 genes to the cytogenetic map and implications for human disorders. Genomics 72: 223-230.

Ühlein, M., Weglöhner, W., Urlauβ, H., and Wittmann-Liebold, B. (1998) Functional

implications of ribosomal protein L2 in protein biosynthesis as shown by *in vivo* replacement studies. Biochem J. 331: 423-430.

Valle, M., Zavialov, A., Sengupta, J., Rawat, U., Ehrenberg, M., and Frank, J. (2003) Locking and unlocking of ribosomal motions. Cell 114: 123-134.

van Beekvelt, C.A., de Graaff-Vincent, M., Faber, A.W., van't Riet, J., Venema, J., and Raué, H.A. (2001) All three functional domains of the large ribosomal subunit protein L25 are required for both early and late pre-rRNA processing steps in *Saccharomyces cerevisiae*. Nuc. Acids Res. 29: 5001-5008.

van Beekvelt, C.A., Kooi, E.A., de Graaff-Vincent, M., van't Riet, J., Venema, J., and Raué, H.A. (2000) Domain III of *Saccharomyces cerevisiae* 25S ribosomal RNA: its role in binding of ribosomal protein L25 and 60S subunit formation. J. Mol. Biol. 296: 7-17.

Van Lijsebettens, M., Vanderhaegen, R., De Block, M., Bauw, G., Villarroel, R., and Van Montagu, M. (1994) An S18 ribosomal protein gene copy at the *Arabidopsis PFL* locus affects plant development by its specific expression in meristems. EMBO J. 13: 3378-3388.

Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E.Jr., Hieter, P., Vogelstein, B., and Kinzler, K.W. (1997) Characterization of the yeast transcriptome. Cell 88: 243-251.

Verschoor, A., Srivastava, S., Grassucci, R., and Frank, J. (1996) Native 3D structure of eukaryotic 80S ribosome: morphological homology with the *E. coli* 70S ribosome. J. Cell Biol. 133: 495-505.

Verschoor, A., Warner, J.R., Srivastava, S., Grassucci, R.A., and Frank, J. (1998) Three-dimensional structure of the yeast ribosome. Nuc. Acids Res. 26: 655-661.

Vester, B., and Garrett, R.A. (1984) Structure of a protein L23-RNA complex located at the A-site domain of the ribosomal peptidyl transferase centre. J. Mol. Biol. 179: 431-452.

Vignais, M.-L., Woudt, L.P., Wassenaar, G.M., Mager, W.H., Sentenac, A., and Planta, R.J. (1987) Specific binding of TUF factor to upstream activation sites of yeast ribosomal protein genes. EMBO J. 6: 1451-1457.

Vilardell, J., Chartrand, P., Singer, R.H., and Warner, J.R. (2000) The odyssey of a

regulated transcript. RNA 6: 1773-1780.

Vishwanath, P., Favaretto, P., Hartman, H., Mohr, S.C., and Smith, T.F. (2004) Ribosomal protein-sequence block structure suggests complex prokaryotic evolution with implications for the origin of eukaryotes. Mol. Phylo. Evol. 33: 615-625.

Volkov, R.A., Panchuk, I.I., and Schöffl, F. (2003) Heat-stress-dependency and developmental modulation of gene expression: the potential of house-keeping genes as internal standards in mRNA expression profiling using real-time RT-PCR. J Expt. Bot. 54: 2343-2349.

Wade, J.T., Hall, D.B., and Struhl, K. (2004) The transcription factor Ifh1 is a key regulator of yeast ribosomal protein genes. Nature 432: 1054-1058.

Wahl, M.C., and Möller, W. (2002) Structure and function of the acidic ribosomal stalk proteins. Curr. Prot. Pept. Sci. 3: 93-106.

Wang, A., and Crowley, D.E. (2005) Global gene expression responses to cadmium toxicity in *Escherichia coli*. J. Bacteriol. 187: 3259-3266.

Warner, J.R. (1989) Synthesis of ribosomes in *Saccharomyces cerevisiae*. Microbiol. Rev. 53: 256-271.

Warner, J.R. (1999) The economics of ribosome biosynthesis in yeast. Trends Bioch. Sci. 24: 437-440.

Warner, J.R., Mitra, G., Schwindinger, W.F., Studeny, M., and Fried, H.M. (1985) *Saccharomyces cerevisiae* coordinates accumulation of yeast ribosomal proteins by modulating mRNA splicing, translational initiation, and protein turnover. Mol. Cell. Biol. 5: 1512-1521.

Watkins, J.F., Sung, P., Prakash, S., and Prakash, L. (1993) The extremely conserved amino terminus of RAD6 ubiquitin-conjugating enzyme is essential for amino-end rule-dependent protein degradation. Genes & Dev 7: 250-261.

Watson, J.D., and Crick, F.H.C. (1953a) A structure for deoxyribose nucleic acid. Nature 171: 737-738.

Watson, J.D., and Crick, F.H.C. (1953b) Genetical implications of the structure of deoxyribonucleic acid. Nature 171: 964-967.

Watson, J.D., and Crick, F.H.C. (1953c) The structure of DNA. Cold Spring Harb.

Symp.Quant. Biol. 18: 123-131.

Watson, J.D. (1964) The synthesis of proteins upon ribosomes. Bull. Soc. Chim. Biol. 46: 1399-1425.

Weigel, D., Alvarez, J., Smyth, D.R., Yanofsky, M.F., and Meyerowitz, E.M. (1992) *LEAFY* controls floral meristem identity in Arabidopsis. Cell 69:843-859.

Weijers, D., Franke-van Dijk, M., Vencken, R.-J., Quint, A., Hooykaas, P., and Offringa, R. (2001) An *Arabidopsis* Minute-like phenotype caused by a semi-dominant mutation in a *RIBOSOMAL PROTEIN S5* gene. Development 128: 4289-4299.

Weinger, J.S., Parnell, K.M., Dorner, S., Green, R., and Strobel, S.A. (2004) Substrate-assisted catalysis of peptide bond formation by the ribosome. Nat. Struct. Mol. Biol. 11: 1101-1106.

Weitzmann, C.J., and Cooperman, B.S. (1990) Reconstitution of *Escherichia coli* 50S ribosomal subunits containing puromycin-modified L23: functional consequences. Biochemistry 29: 3458-3465.

Welfle, H., Stahl, J., and Bielka, H. (1972) Studies on proteins of animal ribosomes. XIII. Enumeration of ribosomal proteins of rat liver. FEBS Lett. 26: 228-232.

Williams, M.E., and Sussex, I.M. (1995) Developmental regulation of ribosomal protein L16 genes in *Arabidopsis thaliana*. Plant J. 8: 65-76.

Williams, A.J., Werner-Fraczek, J., Chang, I.-F., and Bailey-Serres, J. (2003) Regulated phosphorylation of 40S ribosomal protein S6 in root tips of maize. Plant Physiol. 132: 2086-2097.

Williamson, J.R. (2003) After the ribosome structures: how are the subunits assembled? RNA 9: 165-167.

Wimberly, B.T., Brodersen, D.E., Clemons, W.M. Jr., Morgan-Warren, R.J., Carter, A.P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit. Nature 407: 327-339.

Woese, C.R. (2001) Translation: in retrospect and prospect. RNA 7: 1055-1067.

Wool, I.G. (1996) Extraribosomal functions of ribosomal proteins. Trends Biochem. Sci. 21, 164-165.

Wool, I.G., Chan, Y.-L., and Gluck, A. (1995) Structure and evolution of mammalian

ribosomal proteins. Biochem. Cell Biol. 73: 933-947.

Wool, I.G., Correll, C.C., and Chan, Y.-L. (2000) Structure and function of the sarcin ricin domain. In: The Ribosome: Structure, Function, Antibiotics, and Cellular Interactions. Edited by: Garrett, R.A., Douthwaite, S.R., Liljas, A., Matheson, A.T., Moore, P.B., and Noller, H.F. ASM Press, Washington, D.C. pp: 461-473.

Woolhead, C.A., McCormick, P.J., and Johnson, A.E. (2004) Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. Cell 116: 725-736.

Woudt, L.P., Smit, A.B., Mager, W.H., and Planta, R.J. (1986) Conserved sequence elements upstream of the gene encoding yeast ribosomal protein L25 are involved in transcription activation. EMBO J. 5: 1037-1040.

Wu, J., Matsui, E., Yamamoto, K., Nagamura, Y., Kurata, N., Takuji, S., and Minobe, Y. (1995) Genomic organization of 57 ribosomal protein genes in rice (*Oryza sativa* L.) through RFLP mapping. Genome 38: 1189-1200.

Xu, L., He, G.-P., Li, A., and Ro, H.-S. (1994) Molecular characterization of the mouse ribosomal protein S24 multigene family: a uniquely expressed intron-containing gene with cell-specific expression of three alternatively spliced mRNAs. Nuc. Acids Res. 22: 646-655.

Yang, Y., Li, R., Qi, M. (2000) *In vivo* analysis of plant promoters and transcription factors by agroinfiltration of tobacco leaves. Plant J. 22:543-551.

Yates, J.L., Arfsten, A.E., and Nomura, M. (1980) *In vitro* expression of *Escherichia coli* ribosomal protein genes: autogenous inhibition of translation. Proc. Natl. Acad. Sci. U.S.A. 77: 1837-1841.

Young, L.W. (2003) High temperature stress and flowering in *Brassica napus* L. PhD thesis, University of Saskatchewan, Canada.

Yu, Y., Ji, H., Doudna, J.A., and Leary, J.A. (2005) Mass spectrometric analysis of the human 40S ribosomal subunit: native and HCVIRES-bound complexes. Prot. Sci. 14: 1438-1446.

Yu, Z., Wright, S.I., and Bureau, T.E. (2000) *Mutator*-like elements in *Arabidopsis thaliana*: structure, diversity, and evolution. Genetics 156:2019-2031.

Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate,

J.H.D., and Noller, H.F. (2001) Crystal structure of the ribosome at 5.5 Å resolution. Science 292: 883-896.

Zahradka, P., Larson, D.E., and Sells, B.H. (1990) Characterization of a mammalian ribosomal protein gene promoter. Biochem. Cell Biol. 68: 949-956.

Zengel, J.M., Vorozheikina, D., Li, X., and Lindahl, L. (1995) Regulation of the *Escherichia coli* S10 ribosomal protein operon by heterologous L4 ribosomal proteins. Biochem. Cell Biol. 73: 1105-1112.

Zhang, Y., and Berger, S.A. (2004) Increased calcium influx and ribosomal content correlate with resistance to endoplasmic reticulum stress-induced cell death in mutant leukemia cell lines. J. Biol. Chem. 279: 6507-6516.

Zhang, Z., Harrison, P., and Gerstein, M. (2002) Identification and analysis of over 2000 ribosomal protein pseudogenes in the human genome. Genome Res. 12: 1466-1482.

Zurdo, J., González, C., Sanz, J.M., Rico, M., Remacha, M., and Ballesta, J.P.G. (2000) Structural differences between *Saccharomyces cerevisiae* ribosomal stalk proteins P1 and P2 support their functional diversity. Biochemistry 39: 8935-8943.